Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/005080

International filing date: 17 February 2005 (17.02.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/546,075
Filing date: 18 February 2004 (18.02.2004)

Date of receipt at the International Bureau: 23 March 2005 (23.03.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1295760

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

March 14, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/546,075*
FILING DATE: *February 18, 2004*
RELATED PCT APPLICATION NUMBER: *PCT/US05/05080*

Certified by

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

16562 U.S. PTO
021804

Please type a plus sign (+) inside this box → ⊞

PTO/SB/16 (8-00)
Approved for use through 10/31/2002. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CAR 1.53(c).

22896 U.S PTO 60/546075
021804

INVENTOR(S)		
Given name (first and middle [if any])	Family Name or Surname	Residence (city and either State or Foreign Country)
Kambiz Dennis J. Jason M.	Shekdar Sawchuk Montez	New York, New York Hoboken, New Jersey Hoboken, New Jersey

☐ *Additional inventors are being named on the __ separately numbered sheets attached hereto*

TITLE OF THE INVENTION (280 characters max)

ISOLATION OF LIVING CELLS, PREPARATION OF CELL LINES QUANTIFICATION OF PRESELECTED CELLULAR RIBONUCLEIC ACID SEQUENCES BASED ON SIGNALING PROBES

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☒ Customer Number 1473 → *Place Customer Number Bar Code Label here*
OR *Type Customer Number here*

☒ Firm or Individual Name	Fish & Neave				
Address	1251 Avenue of the Americas				
Address					
City	New York	State	New York	ZIP	10020
Country	U.S.A.	Telephone	212-596-9000	Fax	212-596-9090

ENCLOSED APPLICATION PARTS (*check all that apply*)

☒ Specification *Number of Pages* 92 ☐ CD(s), Number
☒ Drawing(s) *Number of Sheets* 45 ☐ Other (specify)
☒ Application Data Sheet. See 37 CFR 1.76

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ Applicant claims small entity status. See 37 CFR 1.27.
☒ A check or money order is enclosed to cover the filing fees
☒ The Commissioner is hereby authorized to charge any underpayment or credit any overpayment of filing fees to Deposit Account Number: 06-1075
☐ Payment by credit card. Form PTO-2038 is attached.

FILING FEE AMOUNT ($): $80.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.
☒ No.
☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____

Respectfully submitted,
SIGNATURE L. Su
TYPED or PRINTED NAME Li Su
TELEPHONE 212-596-9000

Date 02/17/04
REGISTRATION NO. *(if appropriate)* 45,141
Docket Number: CHROMO/003 PROV

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Washington, D.C. 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

Express Mail No. EV079959879US

Docket No. CHROMO/003 PROV

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Kambiz Shekdar et al.	
Application No.	:	Not yet assigned	Confirmation No.:
Filed	:	Concurrently herewith	
For	:	ISOLATION OF LIVING CELLS, PREPARATION OF CELL LINES QUANTIFICATION OF PRESELECTED CELLULAR RIBONUCLEIC ACID SEQUENCES BASED ON SIGNALING PROBES	

"Express Mail" mailing label number: EV079959879US.
Date of Deposit: February 17, 2004 .

I hereby certify that this paper/fee is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 C.F.R. § 1.10 on the date indicated above and is addressed to Mail Stop Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

L. L
Li Su

Encls.:

(1) Provisional Application for Patent Cover Sheet (in duplicate);

(2) Express Mail Certification;

(3) Ninety-two (92) pages of specification;

(4) Forty-five (45) sheets of drawings;

(5) Print EFS Form;

(6) Check in the amount of $80.00; and

(7) Postcard.

ISOLATION OF LIVING CELLS, PREPARATION OF CELL LINES QUANTIFICATION OF PRESELECTED CELLULAR RIBONUCLEIC ACID SEQUENCES BASED ON SIGNALING PROBES

Background of the Invention

[0001] Nucleic acid probes that recognize and report the presence of a specific nucleic acid sequence have been used to detect specific nucleic acids primarily in *in vitro* reactions. See, for example, U.S. Patent 5,925,517, incorporated herein by reference. One type of probe is designed to have a hairpin-shaped structure, with a central stretch of nucleotides complementary to the target sequence, and termini comprising short mutually complementary sequences. See, for example, Tyagi and Kramer, *Nature Biotechnology*, 14, 303-308 (1996), incorporated herein by reference. One terminus of the hairpin-shape designed probe is covalently bound to a fluorophore and the other to a quenching moiety. When in their native state with hybridized termini, the proximity of the fluorophore and the quencher is such that relatively little or essentially no fluorescence is produced. The hairpin-shape designed probe undergoes a spontaneous signaling conformational change when hybridized to its target nucleic acid. Researchers have used the hairpin-shaped probe to perform *in-situ* visualization of messenger RNA (Matsuo, 1998, *Biochim. Biophys. Acta* 1379:178-184) in living cells.

Summary of the Invention

[0002] The present invention provides methods and compositions comprising novel signaling probes for analyzing or isolating cells or generating cell lines

expressing one or more RNA. The method is based on the detection of signal produced by the probes upon their hybridization with target sequence. The RNA may be introduced into the cells via a DNA construct, or the cells may be suspected of expressing the RNA endogenously. The DNA construct may further encode a tag sequence, and the signaling probe is complementary to the tag sequence. In one embodiment, the isolated cells, or generated cell lines are functionally null for expression or have reduced expression of one or more preselected proteins or RNAs. The invention also provides a method of generating transgenic animals using cells that are isolated according to the methods described.

[0003] The invention provides signaling probes used in a method for quantifying the expression level of one or more RNA transcripts. In addition, the signaling probes are used in a method for identifying a compound or RNA sequence that modulates transcription of at least one preselected RNA. In another embodiment, the signaling probes are used in a method for identifying genetic recombinational events in living cells.

[0004] The present invention also provides protease probes. In one embodiment, the signaling or protease probe comprises two separate strands of nucleic acid or modified nucleic acid one or more portions of which anneal to each other, and at least one terminus of one strand is adjacent to a terminus of the other strand. The nucleic acid may be DNA or RNA. For the signaling probe with two separate strands, in one embodiment, one strand has at least a quencher moiety on one terminus, and the other strand has at least a fluorophore on the adjacent terminus. For the protease probe, in one embodiment, one strand has at least a proteolytic enzyme on one terminus, and the other strand has at least an inhibitor of the proteolytic enzyme on the adjacent terminus.

[0005] In another embodiment, the signaling or protease probe is designed to comprise at least a mutually complementary region and at least a non-complementary region. In one embodiment, at least one non-complementary region may be designed to form a loop region. In one embodiment, the probe is designed to form at least a stem-loop structure. In another embodiment, the probe forms a dumbbell structure or a three-arm junction structure. In one embodiment, the signaling probe has at least a fluorophore and at least a quencher moiety at each

terminus of the strand. In one embodiment, the protease probe has a proteolytic enzyme and an inhibitor of the proteolytic enzyme at each terminus of the strand.

[0006] In another embodiment, the signaling or protease probe is chemically modified. One or more of the sugar-phosphodiester type backbone, 2'OH and purine or pyrimidine base is modified. In one embodiment, the deoxyribose backbone is replaced by peptide nucleic acid.

[0007] In one embodiment, the tag sequence is a structural RNA, preferably a three-arm junction structure. In one embodiment, the tag sequence comprises the structure or sequence according to Figure 42 A, B or C. The present invention also provides a DNA construct comprising at least one DNA encoding at least one RNA of interest and the tag sequence. The invention also provides vectors and cells comprising the DNA construct.

[0008] In other embodiments is provided:

1. A method for isolating cells expressing at least one RNA, comprising the steps of:

a) introducing into cells at least a DNA encoding said at least one RNA;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization to said at least one RNA; and

c) isolating said cells that produce the signal.

2. The method of paragraph 1, further comprising the step of generating a cell line or a plurality of cell lines that express said at least one RNA by growing said isolated cells.

3. A method for isolating cells that express at least one of two or more RNAs, comprising the steps of:

a) introducing into cells a first DNA encoding a first RNA;

b) introducing into said cells at least a second DNA encoding at least a second RNA;

c) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization to said first RNA;

d) exposing said cells to at least a second signaling probe that produces a detectable signal upon hybridization to said at least second RNA; and

e) isolating cells that produce at least one of said signals upon hybridization of said signaling probes to their respective RNAs.

4. The method of paragraph 3 further comprising the step of generating a cell line or a plurality of cell lines that express at least one of said two or more RNAs by growing said isolated cells.

5. A method for isolating a plurality of cells, wherein at least a portion of the cells express at least a different RNA, comprising the steps of:

a) introducing into cells a plurality of DNA encoding a plurality of RNA, wherein at least a portion of the cells are introduced at least a different DNA that encodes at least a different RNA;

b) exposing said cells to a plurality of signaling probes sequentially or simultaneously, wherein the signaling probes produce a detectable signal upon hybridization to said plurality of RNA; and

c) isolating said cells that produce the signal.

6. The method of paragraph 5 further comprising the step of generating a plurality of cell lines expressing at least a different RNA by growing said isolated cells.

7. A method for isolating cells expressing at least one RNA, comprising the steps of:

a) introducing into cells at least a DNA encoding said at least one RNA and at least one tag sequence;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization with the tag sequence; and

c) isolating said cells that produce the signal.

8. The method of paragraph 7, further comprising the step of generating a cell line or a plurality of cell lines that express said at least one RNA by growing said isolated cells.

9. A method for isolating cells expressing at least one of two or more RNAs, comprising the steps of:

a) introducing into cells a first DNA encoding a first RNA and at least a first tag sequence;

b) introducing into said cells at least a second DNA encoding at least an additional RNA and at least a second tag sequence, wherein the second tag sequence is the same or different from the first tag sequence;

c) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization with the first tag sequence;

d) exposing said cells to at least a second signaling probe that produces a detectable signal upon hybridization with the second tag sequence; and

e) isolating cells that produce at least one of said signals upon hybridization of said signaling probes to their respective RNAs.

10. The method of paragraph 9 further comprising the step of generating a cell line or a plurality of cell lines that express at least one of said two or more RNAs by growing said isolated cells.

11. The method of paragraph 3 or 9, wherein said steps of said first RNA are performed either simultaneously or sequentially with the corresponding steps of said at least one additional RNA.

12. The method of paragraph 3 or 9, wherein the two or more RNAs or proteins encoded by the two or more RNAs are selected from the group consisting of RNAs or proteins in the same or related biological pathway, RNAs or proteins that act upstream or downstream of each other, RNAs or proteins that have a modulating, activating or repressing function to eath other, RNAs or proteins that are dependent on each other for function or activity, RNAs or proteins that form a complex, proteins from a protein family.

13. A method of isolating cells that overexpress at least one RNA comprising the steps of:

a) introducing into cells at least a first DNA encoding said at least one RNA and at least a first tag sequence; and at least a second DNA encoding said at least one RNA and at least a second tag sequence, wherein the introduction of the first and second DNA construct is performed sequentially or simultaneously, wherein the first and second tag sequences are the same or different;

b) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization with said at least first tag sequence, and to at least a second signaling probe that produces a detectable signal upon hybridization with said at least second tag sequence; and

c) isolating cells that produce at least one of said signals upon hybridization of said signaling probes to their respective RNAs.

14. The method of paragraph 13 further comprising the step of generating a cell line or a plurality of cell lines that overexpress said RNA by growing said isolated cells.

15. The method of any one of paragraphs 3, 9 and 13, wherein the first signaling probe produces a different signal than the signal produced by the second signaling probe.

16. The method of paragraph 13, wherein said steps of said first signaling probe are performed either simultaneously or sequentially with the corresponding steps of said second signaling probe.

17. The method of any one of paragraphs 3, 9 and 13, wherein said first DNA and said second DNA are on the same construct or different constructs.

18. A method for isolating a plurality of cells, wherein at least a portion of the cells express at least a different RNA, comprising the steps of:

a) introducing into cells a plurality of DNA encoding a plurality of RNA and at least one tag sequence, wherein at least a portion of the cells are introduced at least a different DNA that encodes at least a different RNA;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization to said tag sequence; and

c) isolating said cells that produce the signal.

19. The method of paragraph 18, wherein the plurality of RNA form at least an expression library.

20. The method of paragraph 18, further comprising the step of generating a plurality of cell lines expressing at least a different RNA by growing said isolated cells.

21. A method of isolating at least one of two or more RNA expression libraries of cells, comprising the steps of:

a) introducing into cells DNA encoding at least a first RNA expression library and at least a first tag sequence;

b) introducing into cells DNA encoding at least a second RNA expression library and at least a second tag sequence, wherein the second tag sequence is the same or different from the first tag sequence;

c) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization to said at least first tag sequence;

d) exposing said cells to at least a second signaling probe that produces a detectable signal upon hybridization to said at least second tag sequence, wherein the detectable signal from the first signaling probe is the same or different from the detectable signal from the second signaling probe; and

e) isolating said cells that produce at least one of said signal.

22. The method of paragraph 21, further comprising the step of generating at least one of said two or more RNA expression libraries by growing said isolated cells.

23. The method of paragraph 5 or 18, wherein the plurality of RNA or proteins encoded by the plurality of RNAs are selected from the group consisting of RNAs or proteins in the same or related biological pathway, RNAs or proteins that act upstream or downstream of each other, RNAs or proteins that have a modulating, activating or repressing function to eath other, RNAs or proteins that are dependent on each other for function or activity, RNAs or proteins that form a complex, proteins from a protein family.

24. The method of paragraph 18, wherein at least a portion of the plurality of DNA encode the same tag sequence.

25. The method of paragraph 7, 9, 13, 18 or 21, wherein the tag sequence comprises multiple target sequences to be recognized by at least a signaling probe.

26. The method of any one of paragraphs 7, 9, 13, 18 and 21, wherein the tag sequence is a structural RNA.

27. The method of paragraph 25, wherein the tag sequence forms a three-arm junction structure.

28. The method of paragraph 27, wherein the stem region comprises 8-9 basepairs, the first stem-loop region comprises 4-6 basepairs and the second stem-loop region comprises 13-17 basepairs.

29. The method of paragraph 27, wherein the stem regions of the three arms further comprise non-complementary regions.

30. The method of paragraph 27, wherein the stem region and the first stem-loop region both further comprise one mismatch region, and the second stem-loop region further comprises 2-7 mismatch or bulge regions.

31. The method of paragraph 27, wherein the linkage between the stem regions has a total of 8-12 nucleotides.

32. The method of paragraph 26, wherein the tag sequence is selected from the group consisting of the sequences shown in Figure 42A, B and C.

33. The method of paragraph 32, wherein the tag sequence forms the more energetically favorable structures predicted by the sequences shown in Figure 42A, B or C.

34. The method of paragraph 27, wherein the target sequence is the region from all or part of the 3' side of the stem of the first stem-loop region, to the linkage between the first and second stem-loop region, to all or part of the 5' side of the stem of the second stem-loop region.

35. The method of any one of paragraphs 7, 9, 13, 18 and 21, wherein at least a portion of the DNA encode multiple identical tag sequences.

36. The method of paragraph 35, wherein at least a portion of the DNA encode up to 50 identical tag sequences.

37. The method of any one of paragraphs 7, 9, 13, 18 and 21, wherein the DNA encoding said tag sequence is in frame with the DNA encoding said RNA.

38. The method of any one of paragraphs 7, 9, 13, 18 and 21, wherein the DNA encoding said tag sequence is out of frame with the DNA encoding said RNA.

39. A method for isolating cells expressing at least one RNA comprising the steps of:

a) providing cells potentially expressing said at least one RNA;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization with said at least one RNA;

c) isolating said cells that produce the signal.

40. The method of paragraph 39, further comprising the step of generating a cell line or a plurality of cell lines that express said at least one RNA by growing said isolated cells.

41. The method of paragraph 39, wherein said cells further potentially express one or more additional RNA, further comprising the steps of:

a) exposing said cells to at least one additional signaling probe that produces a detectable signal upon hybridization with said at least one additional RNA; and

b) isolating cells that produce the signal.

42. The method of paragraph 41, further comprising the step of generating a cell line or a plurality of cell lines that express said at least one RNA and additional RNA by growing said isolated cells.

43. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, and 21 further comprising the step of adding to the cells a compound that modulates or regulates the expression of said RNA, additional RNA or plurality of RNA prior to step a).

44. The method of paragraph 43, wherein the compound induces the expression of said RNA, additional RNA or plurality of RNA.

45. The method of paragraph 5 or 18, wherein said steps of said plurality of RNA are performed either simultaneously or sequentially.

46. A method for isolating cells expressing at least one exogenous RNA and at least one endogenous RNA, comprising the steps of:

a) introducing into cells DNA encoding said at least one exogenous RNA, wherein said cells potentially express at least one endogenous RNA;

b) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization to said at least one exogenous RNA;

c) exposing said cells to at least a second signaling probe that produces a detectable signal upon hybridization to said at least one endogenous RNA; and

d) isolating said cells that produce at least one of said signals upon hybridization of said signaling probes to their respective RNAs.

47. The method of paragraph 46, further comprising the step of generating a cell line or a plurality of cell lines expressing said at least one exogenous RNA, or said at least one endogenous RNA, or both, by growing said isolated cells.

48. The method of paragraph 46, wherein said steps of said exogenous RNA are performed either simultaneously or sequentially with the corresponding steps of said endogenous RNA.

49. The method of paragraph 46, wherein said second signaling probe produces a different signal than the signal produced by the first signaling probe.

50. The method of paragraph 46, wherein the endogenous RNA and the exogenous RNA or proteins encoded by the endogenous and exogenous RNAs are selected from the group consisting of RNAs or proteins in the same or related biological pathway, RNAs or proteins that act upstream or downstream of each other, RNAs or proteins that have a modulating, activating or repressing function to eath other, RNAs or proteins that are dependent on each other for function or activity, RNAs or proteins that form a complex, proteins from a protein family.

51. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21 and 39, wherein the RNA comprises one or more of a messenger RNA that encodes a protein, an RNA that encodes a peptide, an antisense RNA, a siRNA, a tRNA, a structural RNA, a ribosomal RNA, an hnRNA and an snRNA.

52. The method of paragraph 51, wherein said protein is selected from the group consisting of a cell surface-localized protein, secreted protein and an intracellular protein.

53. A method for isolating cells that overexpress at least a first protein and which are functionally null expressing or reduced in expression for at least a second protein, comprising the steps of:

a) introducing into cells at least a first DNA encoding at least one RNA that encodes said at least first protein, and at least a first tag sequence; and at least a second DNA encoding said at least one RNA and at least a second tag sequence, wherein said first and second tag sequences are different;

b) introducing into cells at least one DNA encoding at least one antisense RNA or siRNA that binds to or interferes with the mRNA transcript of said at least second protein;

c) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization with said at least first tag sequence, and to at least a second signaling probe that produces a detectable signal upon hybridization with said at least second tag sequence;

d) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization to said at least one antisense RNA or siRNA; and

e) isolating cells that produce at least one of said signals upon hybridization of said signaling probes to their respective RNAs.

54. The method of paragraph 53, further comprising the step of generating a cell line or a plurality of cell lines overexpressing at least a first protein and which are functionally null expressing or reduced in expression for at least a second protein.

55. The method of paragraph 53, wherein said steps of said first protein are performed either simultaneously or sequentially with the corresponding steps of said second protein.

56. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53, wherein said DNA is operably linked to a conditional promoter.

57. The method of paragraph 56, wherein the promoter is inducible or repressible, and prior to step (a), a minimal amount of an inducer or a sufficient amount of repressor is added to the cells.

58. The method of paragraph 57, wherein the RNA is antisense RNA or siRNA.

59. The method of paragraph 57, wherein the RNA is lethal or damaging to the cell.

60. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53, further comprising the step of selecting the cells after introducing the DNA into cells but prior to exposing said cells to said signaling probe.

61. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53, wherein at least one DNA further encodes at least one drug resistance marker, and said method further comprises the step of selecting cells resistant to at least one drug to which said marker confers resistance.

62. A method of isolating cells comprising a DNA construct encoding an RNA sequence that is under the control of a tissue specific promotor, comprising the steps of:

a) introducing into cells at least one DNA construct encoding at least a first RNA sequence under the control of a constitutive promotor and encoding at least a second RNA sequence under the control of a tissue specific promotor;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization to said first RNA sequence; and

c) isolating said cells that produce said signal.

63. The method of paragraph 62, further comprising the steps of generating a cell line or a plurality of cell lines that comprises a DNA construct encoding an RNA sequence that is under the control of a tissue specific promotor.

64. The method of paragraph 62, wherein the tissue specific promotor controls expression of a selection marker gene.

65. The method of paragraph 62, wherein the selection marker gene is a drug resistance gene or a detectable protein gene.

66. A method of identifying a compound that activates a tissue specific promotor, comprising the steps of:

a) adding a compound to the cells isolated from paragraph 62;

b) identifying the cells by the selection marker;

c) identifying the compound as a compound that activates the tissue specific promotor.

67. A method of isolating cells comprising a DNA construct encoding at least a test RNA sequence and an RNA sequence that is under the control of a tissue specific promotor, comprising the steps of:

a) introducing into cells at least one DNA construct encoding at least one test RNA sequence under the control of a constitutive promotor, at least a second RNA sequence under the control of a second constitutive promotor that is identical or different, and at least a third RNA sequence under the control of a tissue specific promotor;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization to said second RNA sequence; and

c) isolating said cells that produce said signal.

68. The method of paragraph 67, further comprising the steps of generating a cell line or a plurality of cell lines that comprises a DNA construct encoding at least a test RNA sequence and an RNA sequence that is under the control of a tissue specific promotor.

69. The method of paragraph 67, wherein the test RNA sequence is from an expression library.

70. The method of paragraph 67, wherein the tissue specific promotor controls expression of a selection marker gene.

71. The method of paragraph 67, wherein the selection marker gene is a drug resistance gene or a detectable protein gene.

72. A method of identifying a test RNA sequence that activates a tissue specific promotor, comprising the steps of:

a) identifying the isolated cells of paragraph 67 by the selection marker;

b) identifying the test RNA sequence that activates the tissue specific promotor.

73. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53, further comprising the steps of

i) exposing said isolated cells to a signaling probe that produces a detectable signal upon hybridization to the respective RNA;

ii) determining whether the isolated cells express the respective RNAs; or quantitating the level of the signal to determine the level of expression of the respective RNAs.

74. The isolated cells obtained from the method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53, wherein the cells are applied in a cell-based assay.

75. The isolated cells obtained from any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53, wherein the cells are implantable in an animal.

76. A method for generating a transgenic animal that expresses the RNA according to any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53, comprising carrying out the steps of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39 and 53 utilizing embryonic stem cells or cells that can be implanted in an

animal, determining the viability of said stem cells or cells, and using said viable embryonic stem cells or cells to produce said transgenic animal.

77. The method of paragraphs 1, 3, 5, 7, 9, 13, 18, 21 and 39, wherein the RNA is an antisense RNA or siRNA, and at least one preselected protein in the isolated cells is functionally null or has a reduced expression level as a result of the binding of the antisense RNA, or the interference of the siRNA to mRNA transcripts of said at least one preselected protein.

78. The method of paragraph 77, wherein said preselected protein is an alternatively spliced form of a gene product.

79. A method for generating a transgenic animal, wherein at least one preselected protein is functionally null-expressing or is reduced in expression, comprising carrying out the steps of paragraph 77, utilizing embryonic stem cells or cells that are implantable in an animal, determining the viability of said stem cells or cells, and using said viable embryonic stem cells or cells to produce said transgenic animal.

80. A method for quantifying the level of at least one RNA transcript expression in a biological sample comprising the steps of:

a) exposing said biological sample to a first signaling probe which produces a detectable signal upon hybridization with said RNA transcript;

b) quantitating the level of the signal in said biological sample; and

c) correlating said level of signal with said level of said at least one mRNA transcript.

81. The method of paragraph 80, wherein said biological sample is a cellular sample, a tissue sample or preparations derived thereof.

82. The method of paragraph 80 wherein said RNA transcript is one or more of a messenger RNA that encodes a protein, an RNA that encodes a

peptide, an antisense RNA, a siRNA, a tRNA, a structural RNA, a ribosomal RNA, an hnRNA and an snRNA.

83. The method of paragraph 80, wherein said biological sample is fixed.

84. The method of paragraph 80, wherein the level of at least one second RNA transcript expression is quantified in said biological sample using a second signaling probe which produces a detectable signal upon hybridization to said second RNA transcript.

85. A method of identifying a compound that modulates transcription of at least one preselected RNA, comprising the steps of:

a) adding a compound to cells exogenously or endogenously expressing said preslected RNA;

b) exposing said cells to at least one signaling probe which produces a detectable signal upon hybridization with said at least one preselected RNA;

c) quantitating the level of the signal in said cells;

d) identifying cells that have an increase or decrease in signal compared to the signal of cells with no compound added; and

e) identifying compounds that modulate transcription of said at least one preselected RNA.

86. The method of paragraph 85, wherein cells exogenously expressing said preselected RNA were isolated according to the method of paragraph 1 or 7.

87. The method of paragraph 85, wherein the DNA construct comprises a promotor or operator and encodes a repressor, enhancer, or a sequence that modulates transcription.

88. A method of identifying an RNA sequence that modulates transcription of at least one preselected RNA, comprising the steps of:

a) introducing into cells at least a test RNA sequence that potentially modulates transcription of at least one preselected RNA that is exogenously or endogenously expressed;

b) exposing said cells to at least one signaling probe which produces a detectable signal upon hybridization with said at least one preselected RNA;

c) quantitating the level of the signal in said cells;

d) identifying cells that have an increase or decrease in signal compared to the signal of cells with no test RNA sequence; and

e) identifying a test RNA sequence that modulates transcription of said at least one preselected RNA.

89. The method of paragraph 88, wherein cells exogenously expressing said preselected RNA were isolated according to the method of paragraph 1 or 7.

90. The method of paragraph 88, wherein an expression library of RNA is used as the test RNA sequence.

91. A method for identifying genetic recombinational events in living cells comprising the steps of:

a) exposing a cell to a signaling probe that produces a detectable signal upon hybridization with an RNA sequence selected from the group consisting of that transcribed from a recombined sequence and that transcribed from the non-recombined sequence;

b) detecting said cell expressing said RNA sequence.

92. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39, 53, 62, 67, 73, 80, 85, 88 and 91, wherein the signaling probe comprises two separate strands of nucleic acid or modified nucleic acid that form at least a mutually complementary region.

93. The method of paragraph 92, wherein the two separate strands form a continuous mutually complementary region from 5' to 3' end, and the two strands have the same number of nucleotides.

94. The method of paragraph 92, wherein after mutually complementary regions are formed between the two strands, the 5' end of one strand is offset from the other strand, or the 3' end of that strand is offset from the other strand, or both, wherein the offset is up to 10 nucelotides or modified nucleotides.

95. The method of paragraph 92, wherein the two strands form a mutually complementary region of 5 or 6 continuous basepairs at each end.

96. The method of paragraph 92, wherein the two strands are not identical in sequence.

97. The method of paragraph 92, wherein the strand has more than 30 nucleotides or modified nucleotides.

98. The method of paragraph 92, wherein the nucleic acid is DNA, RNA, or both.

99. The method of paragraph 92, wherein the modified nucleic acid comprises peptide nucleic acid, chemically modified DNA or RNA, or a combination thereof.

100. The method of paragraph 92, wherein the modified nucleic acid is chemically modified in one or more of a sugar group, phosphodiester linkage and base group.

101. The method of paragraph 100, wherein the phosphodiester linkage is substituted with a chemical group selected from the group consisting of --OP(OH)(O)O--, -OP(O^-M^+)(O)O--, -OP(SH)(O)O--, -OP(S^-M^+)(O)O--, --NHP($O)_2$O--, --OC($O)_2$O--, --$OCH_2C(O)_2$ NH--, --$OCH_2C(O)_2$O--, --OP(CH_3)(O)O--, --OP($CH_2C_6H_5$)(O)O--, --P(S)(O)O-- and--OC($O)_2$NH--.

102. The method of paragraph 101, wherein the phosphodiester linkage is substituted with --OP(SH)(O)O--, -OP(S^-M^+)(O)O-- or--P(S)(O)O--.

103. The method of paragraph 100, wherein the 2' position of the chemically modified RNA comprises the chemical group selected from the group consisting of a C_1-C_4 alkoxy, OCH_2-CH=CH_2, OCH_2-CH=CH-CH_3, OCH_2-CH=CH-$(CH_2)_n$CH3 (n=0,1 ...30), halogen, C_1-C_6 alkyl and OCH_3.

104. The method of paragraph 100, wherein the chemically modified RNA comprises a 2'-O-methyl substitution.

105. The method of paragraph 92, wherein the signaling probe comprises at least an interacting pair comprising two chemical groups, and one chemical group is at one terminus of one strand, and the other chemical group is at the adjacent terminus of the other strand.

106. The method of paragraph 92, wherein the signaling probe has two interacting pairs, wherein each end of the probe has one interacting pair at the adjacent terminus of both strands.

107. The method of paragraph 92, wherein the interacting pair is selected from the group consisting of a fluorophore and a quencher, a chemiluminescent label and a quencher or adduct, dye dimer, and FRET donor and acceptor, a proteolytic enzyme and an inhibitor of the proteolytic enzyme or another molecule capable of reversibly inactivating the enzyme.

108. The method of paragraph 92, wherein the interacting pair is a fluorophore and a quencher, and cells that fluoresce are isolated.

109. The method of paragraph 92, wherein the signaling probe comprises at least two fluorophores that are the same or different.

110. The method of paragraph 109, wherein the signaling probe comprises at least two fluorophores that are a FRET donor and acceptor pair, or a harvester and an emitter fluorophore.

111. The method of paragraph 108, wherein the step of isolating said cells that fluoresce is carried out using a fluorescence activated cell sorter, a fluorescence microscope or a fluorometer.

112. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39, 53, 62, 67, 73, 80, 85, 88 and 91, wherein the signaling probe comprises a stem-loop structure.

113. The method of paragraph 112, wherein the stem region forms 4 to 6 continuous basepairs.

114. The method of paragraph 112, wherein the stem-loop structure comprises at least an interactive pair comprising two chemical groups, and one chemical group is at each terminus of the strand, wherein the stem region comprises two mutually complementary regions connected via a non-complementary region, the mutually complementary region adjacent to the interactive pair forms 5 to 6 basepairs, and the mutually complementary region adjacent to the loop region forms 4 to 5 basepairs.

115. The method of paragraph 114, wherein the non-complementary region is a single-stranded loop region or a mismatch region.

116. The method of paragraph 112, wherein the stem-loop structure comprises at least an interactive pair comprising two chemical groups, and one chemical group is at each terminus of the strand, wherein the stem region comprises three mutually complementary regions connected via two non-complementary regions, the first mutually complementary region adjacent to the interactive pair forms 4 to 5 basepairs, the second mutually complementary region forms 2 to 3 basepairs, and the third mutually complementary region adjacent to the loop region forms 2 to 3 basepairs.

117. The method of paragraph 116, wherein the non-complementary regions are one or more of a single-stranded loop region and a mismatch region.

118. The method of any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39, 53, 62, 67, 73, 80, 85, 88 and 91, wherein the signaling probe comprises a dumbbell structure.

119. The method of paragraph 118, wherein the dumbbell structure comprises one stem region of 4 continuous basepairs, and one stem region of 3 continuous basepairs.

120. The method of paragraph 118, wherein the two stem regions are connected by a phosphodiester linkage or modified phosphodiester linkage via one arm of the stem regions.

121. The method of paragraph 118, wherein the two stem regions are connected by 1 or 2 nucleotides or modified nucleotides via one arm of the stem regions.

122. The method of paragraphs any one of paragraphs 1, 3, 5, 7, 9, 13, 18, 21, 39, 53, 62, 67, 73, 80, 85, 88 and 91, wherein the signaling probe comprises a three-arm junction structure.

123. The method of paragraph 122, wherein the three-arm junction structure comprises at least an interactive pair comprising two chemical groups, and one chemical group is at each terminus of the strand, wherein the stem region adjacent to the interactive pair forms 3 to 4 continuous basepairs, the stem region of the first stem-loop structure forms 4 to 5 continuous basepairs, and the stem region of the second stem-loop structure forms 2 to 3 continuous basepairs.

124. The method of paragraph 122, wherein the three regions are connected by a phosphodiester linkage or modified phosphodiester linkage via the arms of the stem regions.

125. The method of paragraph 122, wherein the three regions are connected by 1 or 2 nucleotides or modified nucleotides via the arms of the stem regions.

126. The method of any one of paragraphs 112, 118 and 122, wherein the stem-loop structure, dumbbell structure or three-arm junction structure has more than 30 nucleotides or modified nucleotides.

127. The method of any one of paragraphs 112, 118 and 122 , wherein the structure is DNA, RNA, peptide nucleic acid, chemically modified DNA, RNA, or a combination thereof.

128. The method of of any one of paragraphs 112, 118 and 122, wherein the structure is chemically modified in one or more of a sugar group, phosphodiester linkage and base.

129. The method of paragraph 128, wherein the phosphodiester linkage is substituted with the chemical group selected from the group consisting of --OP(OH)(O)O--, -OP(O^-M^+)(O)O--, -OP(SH)(O)O--, -OP(S^-M^+)(O)O--, --NHP($O)_2$O--, --OC($O)_2$O--, --OCH_2C($O)_2$ NH--, --OCH_2C($O)_2$O--, --OP(CH_3)(O)O--, --OP($CH_2C_6H_5$)(O)O--, --P(S)(O)O-- and --OC($O)_2$NH--.

130. The method of paragraph 129, wherein the phosphodiester linkage is substituted with --OP(SH)(O)O--, -OP(S^-M^+)(O)O-- or--P(S)(O)O--.

131. The method of paragraph 128, wherein the 2' position of the chemically modified RNA comprises the chemical group selected from the group consisting of a C_1-C_4 alkoxy, OCH_2-CH=CH_2, OCH_2-CH=CH-CH_3, OCH_2-CH=CH-($CH_2)_n$CH3 (n=0,1 ...30), halogen, or C_1-C_6 alkyl and OCH_3.

132. The method of paragraph 128, wherein the chemically modified RNA has a 2'-O-methyl substitution.

133. The method of any one of paragraphs 112, 118 and 122, wherein the interacting pair is selected from the group consisting of a fluorophore and a quencher, a chemiluminescent label and a quencher or adduct, dye dimer, and a FRET donor and acceptor, a proteolytic enzyme and an inhibitor of the proteolytic enzyme or another molecule capable of reversibly inactivating the enzyme.

134. The method of paragraph 133, wherein the interacting pair is a fluorophore and a quencher, and cells that fluoresce are isolated.

135. The method of paragraph 134, wherein the step of isolating said cells that fluoresce is carried out using a fluorescence activated cell sorter, fluorescence microscope or fluorometer.

136. The method of any one of paragraphs 112, 118 and 122, wherein the signaling probe comprises at least two fluorophores on one terminus of the strand, and a quencher on the other terminus of the strand, wherein the two fluorophores are a FRET donor and acceptor pair.

137. A probe comprising a nucleic acid or modified nucleic acid comprising sequence complementary to a target sequence and mutually complementary sequences, and at least a proteolytic enzyme and at least an inhibitor of the proteolytic enzyme, wherein said probe having, under assay conditions in the absence of said target sequence, a characteristic proteolytic activity whose level is a function of the degree of interaction of said proteolytic enzyme and inhibitor thereof; and wherein under conditions in the presence of an excess of said target sequence, hybridization of the target complement sequence to the target sequence increases the level of said characteristic proteolytic activity.

138. The probe of paragraph 137, wherein the proteolytic enzyme is a site-specific or target-specific protease.

139. The probe of paragraph 137, wherein said proteolytic enzyme inhibitor is a peptide or compound.

140. The probe of paragraph 137, wherein said proteolytic enzyme and said inhibitor of said proteolytic enzyme is selected from the group consisting of aminopeptidase and amastatin, trypsin-like cysteine proteases and antipain, aminopeptidase and bestatin, chymotrypsin like cysteine proteases and chymostatin, aminopeptidase and diprotin A or B, carboxypeptidase A and EDTA, elastase-like serine proteases and elastinal, and thermolysin or aminopeptidase M and 1,10-phenanthroline.

141. The probe of paragraph 137 that comprises two separate strands of nucleic acid or modified nucleic acid that form at least a mutually complementary region.

142. The probe of paragraph 141, wherein the proteolytic enzyme is at one terminus of one strand, and the proteolytic enzyme inhibitor is at the adjacent terminus of the other strand.

143. The probe of paragraph 141, wherein the two separate strands form a continuous mutually complementary region from 5' to 3' end, and the two strands have the same number of nucleotides.

144. The probe of paragraph 141, wherein after mutually complementary regions are formed between the two strands, the 5' end of one strand is offset from the other strand, or the 3' end of that strand is offset from the other strand, or both, wherein the offset is up to 10 nucelotides or modified nucleotides.

145. The probe of paragraph 141, wherein the two strands form a mutually complementary region of 5 or 6 continuous basepairs at each end.

146. The probe of paragraph 145, wherein the two strands are not identical in sequence.

147. The probe of paragraph 141, wherein the strand has more than 30 nucleotides or modified nucleotides.

148. The probe of paragraph 141, wherein the nucleic acid is DNA, RNA, or both.

149. The probe of paragraph 141, wherein the modified nucleic acid comprises peptide nucleic acid, chemically modified DNA or RNA, or a combination thereof.

150. The probe of paragraph 141, wherein the modified nucleic acid is chemically modified in one or more of a sugar group, phosphodiester linkage and base group.

151. The probe of paragraph 150, wherein the phosphodiester linkage is substituted with a chemical group selected from the group consisting of --OP(OH)(O)O--, -OP(O^-M^+)(O)O--, -OP(SH)(O)O--, -OP(S^-M^+)(O)O--, --NHP(O_2)O--, --OC(O_2)O--, --$OCH_2C(O)_2$ NH--, --$OCH_2C(O)_2O$--, --OP(CH_3)(O)O--, --OP($CH_2C_6H_5$)(O)O--, --P(S)(O)O-- and--OC(O_2)NH--.

152. The probe of paragraph 150, wherein the phosphodiester linkage is substituted with --OP(SH)(O)O--, -OP(S^-M^+)(O)O-- or--P(S)(O)O--.

153. The probe of paragraph 150, wherein the 2' position of the chemically modified RNA comprises the chemical group selected from the group consisting of a C_1-C_4 alkoxy, OCH_2-CH=CH_2, OCH_2-CH=CH-CH_3, OCH_2-CH=CH-$(CH_2)_n$CH3 (n=0,1 ...30), halogen, C_1-C_6 alkyl and OCH_3.

154. The probe of paragraph 150, wherein the chemically modified RNA comprises a 2'-O-methyl substitution.

155. The probe of paragraph 137 that comprises a stem-loop structure.

156. The probe of paragraph 155, wherein the stem region forms 5 to 6 continuous basepairs.

157. The probe of paragraph 155, wherein the stem-loop structure comprises at least a proteolytic enzyme and an inhibitor of said proteolytic enzyme at each terminus of the strand, wherein the stem region comprises two mutually complementary regions connected via a non-complementary region, the mutually complementary region adjacent to the terminus of the strand forms 5 to 6 basepairs, and the mutually complementary region adjacent to the loop region forms 4 to 5 basepairs.

158. The probe of paragraph 157, wherein the non-complementary region is a single-stranded loop region or a mismatch region.

159. The probe of paragraph 155, wherein the stem-loop structure comprises at least a proteolytic enzyme and an inhibitor of said proteolytic enzyme at each terminus of the strand, wherein the stem region comprises three mutually complementary regions connected via two non-complementary regions, the first mutually complementary region adjacent to the terminus of the strand forms 4 to 6 basepairs, the second mutually complementary region forms 2 to 3 basepairs, and the third mutually complementary region adjacent to the loop region forms 2 to 3 basepairs.

160. The probe of paragraph 159, wherein the non-complementary regions are one or more of a single-stranded loop region and a mismatch region.

161. The probe of paragraph 137 that comprises a dumbbell structure.

162. The probe of paragraph 161, wherein the dumbbell structure comprises one stem region of 4 continuous basepairs, and one stem region of 3 continuous basepairs.

163. The probe of paragraph 161, wherein the two stem regions are connected by a phosphodiester linkage or modified phosphodiester linkage via one arm of the stem regions.

164. The probe of paragraph 161, wherein the two stem regions are connected by 1 or 2 nucleotides or modified nucleotides via one arm of the stem regions.

165. The probe of paragraph 137 that comprises a three-arm junction structure.

166. The probe of paragraph 165, wherein the three-arm junction structure comprises at least a proteolytic enzyme and an inhibitor of said

proteolytic enzyme at each terminus of the strand, wherein the stem region adjacent to the terminus of the strand forms 3 to 4 continuous basepairs, the stem region of the first stem-loop structure forms 4 to 5 continuous basepairs, and the stem region of the second stem-loop structure forms 2 to 3 continuous basepairs.

167. The probe of paragraph 165, wherein the three regions are connected by a phosphodiester linkage or modified phosphodiester linkage via the arms of the stem regions.

168. The probe of paragraph 165, wherein the three regions are connected by 1 or 2 nucleotides or modified nucleotides via the arms of the stem regions.

169. The probe of paragraph 155, 161 or 165, wherein the stem-loop structure, dumbbell structure or three-arm junction structure has more than 30 nucleotides or modified nucleotides.

170. The probe of paragraph 155, 161 or 165, wherein the structure is DNA, RNA, peptide nucleic acid, chemically modified DNA, RNA, or a combination thereof.

171. The probe of paragraph 155, 161 or 165, wherein the structure is chemically modified in one or more of a sugar group, phosphodiester linkage and base.

172. The probe of paragraph 171, wherein the phosphodiester linkage is substituted with the chemical group selected from the group consisting of --OP(OH)(O)O--, -OP(O^-M^+)(O)O--, -OP(SH)(O)O--, -OP(S^-M^+)(O)O--, --NHP(O_2)O--, --OC(O_2)O--, --$OCH_2C(O)_2$ NH--, --$OCH_2C(O)_2$O--, --OP(CH_3)(O)O--, --OP($CH_2C_6H_5$)(O)O--, --P(S)(O)O-- and --OC(O_2)NH--.

173. The probe of paragraph 171, wherein the phosphodiester linkage is substituted with --OP(SH)(O)O--, -OP(S^-M^+)(O)O-- or--P(S)(O)O--.

174. The probe of paragraph 171, wherein the 2' position of the chemically modified RNA comprises the chemical group selected from the group

consisting of a C_1-C_4 alkoxy, OCH_2-$CH=CH_2$, OCH_2-$CH=CH$-CH_3, OCH_2-$CH=CH$-$(CH_2)_nCH3$ (n=0,1 ...30), halogen, or C_1-C_6 alkyl and OCH_3.

175. The probe of paragraph 171, wherein the chemically modified RNA has a 2'-O-methyl substitution.

176. A DNA construct comprising at least one DNA encoding at least one RNA of interest and a tag sequence, wherein the tag sequence forms a three-arm junction structure, and the stem region comprises 8-9 basepairs, the first stem-loop region comprises 4-6 basepairs and the second stem-loop region comprises 13-17 basepairs.

177. The DNA construct of paragraph 176, wherein the stem regions of the three arms further comprise non-complementary regions.

178. The DNA construct of paragraph 176, wherein the stem region and the first stem-loop region both further comprise one mismatch region, and the second stem-loop region further comprises 2-7 mismatch or bulge regions.

179. The DNA construct of paragraph 176, wherein the linkage between the stem regions has a total of 8-12 nucleotides.

180. A DNA construct comprising at least one DNA encoding at least one RNA of interest and a tag sequence, wherein the tag sequence is selected from the group consisting of the sequences shown in Figure 42A, B and C.

181. A DNA construct comprising at least one DNA encoding at least one RNA of interest and a tag sequence, wherein the tag sequence forms the more energetically favorable structures predicted by the sequences shown in Figure 42A, B or C.

182. A vector comprising the DNA construct of any one of paragraphs 176 to 181.

183. A cell comprising the vector of paragraph 182 or the DNA construct of any one of paragraphs 176 to 181.

184. The cell of paragraph 183, that is selected from the group consisting of immortalized, primary, stem and germ cell.

185. The cell of paragraph 183 that is selected from the group consisting of HeLa cell, NIH3T3 cell, HEK293 cell and CHO cell.

186. A library of stable mammalian cell lines comprising at least 10,000 cell lines each comprising at least one stably integrated expressed sequence.

187. A library of stable mammalian cell lines comprising at least 500 cell lines each comprising at least two stably integrated sequences.

188. A library of stable mammalian cell lines comprising at least 50 cell lines each comprising at least three stably integrated sequences.

189. A library of stable mammalian cell lines comprising at least 20 cell lines each comprising at least four stably integrated sequences.

190. The library of any one of paragraphs 186-189, wherein the cell lines further comprise a drug resistance gene.

191. A library of stable mammalian cell lines comprising at least 50 cell lines each comprising at least one stably integrated sequence, wherein the cell lines lack a drug resistance gene.

192. A library of stable mammalian cell lines comprising at least 20 cell lines each comprising at least two stably integrated sequences, wherein the cell lines lack a drug resistance gene.

193. The library of any one of paragraphs 186-192, wherein each cell line comprises a variable library sequence.

194. The library of paragraph 193, wherein the variable sequence of said expression library is selected from the group consisting of genomic, genomic untranslated, genomic translated, gene, cDNA, EST, oligo, random, RNA, protein, protein domain, peptide, intronic, exonic, tag, or linker sequence, or

combination thereof or recombination thereof, or one or more of the unmodified, mutagenized, randomized, shuffled or recombined sequences.

195. The library of any one of paragraphs 186-192, wherein the library was generated using at least a pool or mixture of genetic sequences having unknown sequence identity.

196. The library of paragraph 195, wherein said sequence having unknown identity has shared sequence homology, functional significance, or related origin.

197. The library of any one of paragraphs 186-192, wherein the library is a collection of individually synthesized constructs comprising specific sequences having known identities.

198. The library of paragraph 197, wherein said specific sequences have shared sequence homology, functional significance, or related origin.

199. The library of any one of paragraphs 186-192, wherein said expressed sequence is under the control of a constitutive promotor.

200. The library of any one of paragraphs 186-192, wherein said expressed sequence is under the control of a conditional promotor selected from the group consisting of an inducible, repressible, tissue-specific, temporal or heat-shock promotor.

201. The library of any one of paragraphs 186-192, wherein the library is used in a cell-based screening assay.

202. The library of paragraph 201, wherein the cell-based screening assay is carried out in parallel for all cell lines in said library.

203. The library of paragraph 201, wherein the cell-based screening assay is carried out for a portion of the cell lines in said library.

204. A nucleic acid or modified nucleic acid molecule comprising the sequence of any one of FP1 to FP18 according to Figures 8 to 24 and 41.

205. The nucleic acid or modified nucleic acid molecule of paragraph 204 further comprising an interactive pair.

206. The nucleic acid or modified nucleic acid molecule of paragraph 205, wherein the interactive pair is selected from the group consisting of a fluorophore and a quencher, a chemiluminescent label and a quencher or adduct, dye dimer, and a FRET donor and acceptor, a proteolytic enzyme and an inhibitor of the proteolytic enzyme or another molecule capable of reversibly inactivating the enzyme.

207. A nucleic acid or modified nucleic acid molecule that hybridizes to the sequences according to Figure 42 D1, 2 or 3.

208. The nucleic acid or modified nucleic acid molecule of paragraph 207 further comprising an interactive pair.

209. The nucleic acid or modified nucleic acid molecule of paragraph 208, wherein the interactive pair is selected from the group consisting of a fluorophore and a quencher, a chemiluminescent label and a quencher or adduct, dye dimer, and a FRET donor and acceptor, a proteolytic enzyme and an inhibitor of the proteolytic enzyme or another molecule capable of reversibly inactivating the enzyme.

210. A nucleic acid or modified nucleic acid molecule comprising the sequence in Figure 42 A, B or C.

211. A nucleic acid or modified nucleic acid molecule comprising the sequence in Figure 42 D1 or D2.

[0009] These and other aspects of the invention will be appreciated from the following detailed description.

Brief Description of the Drawings

[0010] Figure 1 and 2 depict signaling or protease probes with two separate strands. The interacting chemical groups are shown as ovals. The different ovals present one embodiment of the invention, wherein the dark oval indicates a quencher moiety, and the white and grey ovals indicate different fluorophores.

[0011] Figure 3, 4 and 5 depict signaling or protease probes designed to have a stem-loop structure. The interacting chemical groups are shown as ovals. The different ovals present one embodiment of the invention, wherein the dark oval indicates a quencher moiety, and the white oval indicates a fluorophore.

[0012] Figure 6 depicts signaling or protease probes with a three-arm junction structure. The interacting chemical groups are shown as ovals. The different ovals present one embodiment of the invention, wherein the dark oval indicates a quencher moiety, and the white oval indicates a fluorophore.

[0013] Figure 7 depicts signaling or protease probes with a dumbbell structure. The interacting chemical groups are shown as ovals. The different ovals present one embodiment of the invention, wherein the dark oval indicates a quencher moiety, and the white oval indicates a fluorophore.

[0014] Figure 8 shows the sequence and predicted native conformation of fluorescent probe FP1. The FP1 sequence comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. Panel A shows the predicted structure of the sequence using DNA folding programs according to *Nucleic Acids Res.* 31: 3429-3431 (2003). Panel B shows predicted self dimerization of the FP1 sequence according to the oligoanalyzer 3.0 software available at http://biotools.idtdna.com/analyzer/oligocalc.asp. In both Panels A and B, the flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0015] Figure 9 shows the sequence and predicted native conformation of fluorescent probe FP2. The sequence comprises bases which are designed to be

complementary to the target sequence and additional flanking bases. The flanking bases are underlined. Panel A shows the predicted structure of the sequence using DNA folding programs according to *Nucleic Acids Res.* 31: 3429-3431 (2003). It is likely that all or part of the shaded region form Watson-Crick basepairs, thereby forming a three-arm junction. Panel B shows predicted self dimerization of the FP2 sequence according to the oligoanalyzer 3.0 software available at http://biotools.idtdna.com/analyzer/oligocalc.asp. In both Panels A and B, the flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0016] Figure 10 shows the sequence and predicted native conformation of fluorescent probe FP3. The FP3 sequence comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. The figure shows predicted self dimerization of the FP3 sequence according to the oligoanalyzer 3.0 software available at http://biotools.idtdna.com/analyzer/oligocalc.asp. The flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0017] Figure 11 shows the sequence and predicted native conformation of fluorescent probe FP4. The FP4 sequence comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. The figure shows the predicted structure of the sequence using DNA folding programs according to *Nucleic Acids Res.* 31: 3429-3431 (2003). The flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0018] Figures 12 through 13 show the sequences of fluorescent probes FP5 to FP6. The sequences comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined.

[0019] Figure 14 shows the sequence and predicted native conformation of fluorescent probe FP7. The FP7 sequence comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. The predicted self dimerization of the FP7 sequence

according to the oligoanalyzer 3.0 software available at http://biotools.idtdna.com/analyzer/oligocalc.asp is shown. The flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0020] Figure 15 shows the sequence and predicted native conformation of fluorescent probe FP8. The FP8 sequence comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. Panel A shows the predicted structure of the sequence using DNA folding programs according to *Nucleic Acids Res.* 31: 3429-3431 (2003). Panel B shows predicted self dimerization of the FP1 sequence according to the oligoanalyzer 3.0 software available at http://biotools.idtdna.com/analyzer/oligocalc.asp. In both Panels A and B, the flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0021] Figures 16 through 20 show the sequences of fluorescent probes FP9 to FP13. The sequences comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined.

[0022] Figure 21 shows the sequence and predicted native conformation of fluorescent probe FP14. The FP14 sequence comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. The figure shows the predicted structure of the sequence using DNA folding programs according to *Nucleic Acids Res.* 31: 3429-3431 (2003). The flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0023] Figures 22 through 24 show the sequence and predicted native conformation of fluorescent probes FP15 to 17, respectively. The sequences comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. Panel A shows the predicted structure of the sequence using DNA folding programs according to *Nucleic Acids Res.* 31: 3429-3431 (2003). Panel B shows predicted self dimerization of the FP15 sequence according to the oligoanalyzer 3.0 software

available at http://biotools.idtdna.com/analyzer/oligocalc.asp. In both Panels A and B, the flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0024] Figure 25 shows the number of cells observed in reference to the fluorescence intensity during the FACS process. Figure 25A shows the profile for control NIH3T3 cells exposed to signaling probe FP1. The fluorophore Alexafluor633 was used in FP1. Control NIH3T3 cells were untransfected with plasmid and do not contain target sequence. Figure 25B shows the profile for transfected NIH3T3 cells exposed to signaling probe FP1. The transfected NIH3T3 cells contained DNA encoding the RNA of interest and a tag1 sequence shown in Figure 42A. Figure 25C is an overlay of Figure 25A and 25B. Figure 25C shows that FACS distinguishes cells transfected with plasmid encoding target sequences from untransfected control cells.

[0025] Figure 26 shows the number of cells observed in reference to the fluorescence intensity during the FACS process. Figure 26A shows the profile for control NIH3T3 cells exposed to signaling probe FP2. The fluorophore Alexafluor680 was used in FP2. Control NIH3T3 cells were untransfected with plasmid and do not contain target sequence. Figure 26B shows the profile for transfected NIH3T3 cells exposed to signaling probe FP2. The transfected NIH3T3 cells contained DNA encoding the RNA of interest and a tag2 sequence shown in Figure 42B. Figure 26C is an overlay of Figure 26A and 26B. Figure 26C shows that FACS distinguishes cells transfected with plasmid encoding target sequences from untransfected control cells.

[0026] Figure 27 shows the number of cells observed in reference to the fluorescence intensity during the FACS process. Figure 27A shows the profile for control NIH3T3 cells exposed to signaling probe FP3. The fluorophore fluorescein was used in FP3. Control NIH3T3 cells were untransfected with plasmid and do not contain target sequence. Figure 27B shows the profile for transfected NIH3T3 cells exposed to signaling probe FP3. The transfected NIH3T3 cells contained DNA encoding the RNA of interest and a tag3 sequence shown in Figure 42C. Figure 27C is an overlay of Figure 27A and 27B. Figure 27C shows that FACS

distinguishes cells transfected with plasmid encoding target sequences from untransfected control cells.

[0027] Figure 28 shows the number of cells observed in reference to the fluorescence intensity during the FACS process. Figure 28A shows the profile for control HeLa cells exposed to signaling probe FP1. The fluorophore fluorescein was used in FP1. Control HeLa cells were untransfected with plasmid and do not contain target sequence. Figure 28B shows the profile for transfected HeLa cells exposed to signaling probe FP1. The transfected HeLa cells contained DNA encoding the reverse vav RNA. Figure 28C is an overlay of Figure 28A and 28B. Figure 28C shows that FACS distinguishes cells transfected with plasmid encoding target sequences from untransfected control cells.

[0028] Figure 29 shows the number of cells observed in reference to the fluorescence intensity during the FACS process. Figure 29A shows the profile for control HeLa cells exposed to signaling probe FP8. The fluorophore fluorescein was used in FP8. Control HeLa cells were untransfected with plasmid and do not contain target sequence. Figure 29B shows the profile for transfected HeLa cells exposed to signaling probe FP8. The transfected HeLa cells contained DNA encoding the reverse vav RNA. Figure 29C is an overlay of Figure 29A and 29B. Figure 29C shows that FACS distinguishes cells transfected with plasmid encoding target sequences from untransfected control cells.

[0029] Figure 30 shows the number of cells observed in reference to the fluorescence intensity during the FACS process. Figure 30A shows the profile for control HeLa cells exposed to signaling probe FP5. The fluorophore fluorescein was used in FP5. Control HeLa cells were untransfected with plasmid and do not contain target sequence. Figure 30B shows the profile for transfected HeLa cells exposed to signaling probe FP5. The transfected HeLa cells contained DNA encoding the reverse vav RNA. Figure 30C is an overlay of Figure 30A and 30B. Figure 30C shows that FACS distinguishes cells transfected with plasmid encoding target sequences from untransfected control cells.

[0030] Figure 31 shows the number of cells observed in reference to the fluorescence intensity during the FACS process. Figure 31A shows the profile for control HeLa cells exposed to signaling probe FP9. The fluorophore fluorescein

was used in FP9. Control HeLa cells were untransfected with plasmid and do not contain target sequence. Figure 31B shows the profile for transfected HeLa cells exposed to signaling probe FP9. The transfected HeLa cells contained DNA encoding the reverse vav RNA. Figure 31C is an overlay of Figure 31A and 31B. Figure 31C shows that FACS distinguishes cells transfected with plasmid encoding target sequences from untransfected control cells.

[0031] Figure 32 shows fluorescence images of drug selected HeLa cells transfected with an expression plasmid encoding a portion of the sequence of vav cloned in reverse orientation (referred to r-vav) as well as a drug resistance gene. The cells were exposed to fluroescent probes (FP) designed to recognize the same target sequence within r-vav (5' GTTCTTAAGGCACAGGAACTGGGA 3'). The images were obtained using a fluorescence microscope and filters designed to detect fluorescence from Fam. All FPs used here were labeled using FAM except FP1, which was labeled using fluorescein. Panel A, B, C, D each were exposed to FP10, FP11, FP12 and FP13.

[0032] Figure 33 shows the fluorescence image of cells transfected with constructs encoding RNA and tag sequences designed to be recognized by FP15 (Panels A, B, D) or transfected with control constructs encoding the same RNA but not the tag sequence (Panel C or E). The tag sequences used were termed 6-5, 6-7, or 6B10, which contained 1, 2 and 3 copies of the target sequence, respectively, which FP15 was designed to recognize. Cells transfected with constructs comprising tag sequences designed to be recognized by FP15 (Panels A, B and D) exhibited greater fluorescence than control cells (Panel C or E).

[0033] Figure 34 shows the fluorescence image of cells transfected with constructs encoding RNA and tag sequence 6CA4 designed to be recognized by FP16 (Panel A) or transfected with control constructs encoding the same RNA but not the tag sequence (Panel B). Cells transfected with constructs comprising the tag sequence (Panel A) exhibited greater fluorescence than control cells (Panel B).

[0034] Figure 35A shows a portion of (underlined) the reverse complement of the vav DNA sequence (r-vav DNA) selected for forming the tag sequence. This sequence was cloned into an expression plasmid designed to express r-vav mRNA. The sequence indicated in bold is the target sequence for certain fluorescent

probes. Figure 35B shows the sequences underlined in Figure 35A after they have been combined to form a tag sequence (tag1 sequence).

[0035] Figure 36A shows the predicted structure of part of the r-vav RNA using RNA folding programs in *Nucleic Acids Res.* 31: 3429-3431 (2003). Figure 36B is the predicted structure of the tag1 sequence shown in Figure 35B. The shading indicates the target sequence designed to be recognized by some of the fluorescent probes.

[0036] Figures 37A, B and C show the predicted structures for tag 1, 2 and 3 sequence as described in Figure 42. These structures resemble each other but present a different sequence for recognition by fluorescent probes. The prediction was generated using RNA folding programs in *Nucleic Acids Res.* 31: 3429-3431 (2003).

[0037] Figure 38 shows fluorescence signal emitted from FPs in solution in the presence of target or control oligo sequence. Samples were illuminated by UV and photographed. All FPs used here incorporated Fluorescein. Tubes each contained 16ul total consisting of 5ul of a 20uM FP stock, 1.5ul 25mM MgCl2, 8ul 20uM oligo, and 1.5ul of water, having a final magnesium concentration of approximately 2.34 mM. FP1 and FP18 were used here and were synthesized incorporating sulfur linkages between the bases of the sequence designed to recognize target oligos TO-FP1 and TO-FP18, respectively. FP1 is directed against the sequence of target oligo 1 (TO-FP1 5'GTTCTTAAGGCACAGGAACTGGGA3'), and FP 18 is directed against the sequence of target oligo FP18 (TO-FP18 5' TCCCAGTTCCTGTGCCTTAAGAAC3'). The sequences of TO-FP1 and TO-FP18 were reverse complements of each other. TO-FP18 has sequence not targeted by FP1 and served as a control oligo for FP1. TO-FP1 has sequence not targeted by FP18 and served as a control oligo for FP18.

In all Panels, the compositions of the tubes are as indicated below:

tube	FP	oligo
1	FP18	TO-FP18
2	FP18	TO-FP1
3	FP1	TO-FP18

4	FP1	TO-FP1

This figure shows that each of the FPs tested were specifically reporting the presence of target sequences by emitting a greater signal in tubes containing oligos having targeted sequence as compared to control tubes containing oligos having non-targeted sequence. Tubes containing FPs in the presence of oligos comprising target sequence are indicated by asterisk.

[0038] Figure 39 shows fluorescence signal emitted from FPs in solution in the presence of target or control oligo sequence. Samples were illuminated by UV and photographed. All FPs used here incorporated FAM. Panels A, B, C each show four tubes to which the same FP was added. Each tube contains a total of 10ul containing 2ul of a 20 uM FP stock and 1ul of a 100um oligo stock in PBS supplemented to 4mM $MgCl_2$. In each Panel, tube 1 contained no oligo stock and instead contained 1ul water, tube 2 contained oligo TO-M1, tube 3 contained oligo TO-M2 and tube 4 contained oligo TO-M3.

[0039] The FPs tested in each Panel are listed below, alongside the oligo which includes sequence that is designed to be recognized by the FP:

Panel	FP	oligo comprising target sequence
A	FP4	TO-M3 (tube 4)
B	FP6	TO-M2 (tube 3)
C	FP7	TO-M3 (tube 4)

[0040] This figure shows that each of the FPs tested were specifically reporting the presence of target sequences by emitting a greater signal in tubes containing oligo having targeted sequence as compared to control tubes containing no oligo or oligo having non-targeted sequence. Tubes containing FPs in the presence of oligos comprising target sequence are indicated by asterisk.

[0041] Sequences from 5' to 3' direction for TO-M1, TO-M2 and TO-M3 are listed below:

TO-M1:

TTTCTCTGTGATCCGGTACAGTCCTTCTGCGCAGGTGGACAGGAA
GGTTCTAATGTTCTTAAGGCACAGGAACTGGGACATCTGGGCCCG
GAAAGCCTTTTTCTCTGTGATCCGGTACAGTCCTTCTGCGCAGGT
GGACAGGAAGGTTCTAATGTTCTT

TO-M2:

TTTAACTGATGGATGGAACAGTCCTTCTGCGCAGGTGGACAGCTT
GGTTCTAATGAAGTTAACCCTGTCGTTCTGCGACATCTGGGCCCG
GAAAGCGTTTAACTGATGGATGGAACAGTCCTTCTGCGCAGGTGG
ACAGCTTGGTTCTAATGAAGTT

TO-M3:

GTAAAGTCAGACATCCGGTACAGTCCTTCTGCGCAGGTGGACAGG
AAGGTTCTAATGTTCTATAGGGTCTGCTTGTCGCTCATCTGGGCC
CGGAGATGCGTAAAGTCAGACATCCGGTACAGTCCTTCTGCGCAG
GTGGACAGGAAGGTTCTAATGTTCTAT

[0042] Figure 40 shows fluorescence signal emitted from FPs in solution in the presence of target or control oligo sequence. Samples were illuminated by UV and photographed. All FPs used here incorporated FAM. FPs 1, 2 and 3 were tested in Panels A, B and C, respectively, and are each designed to recognize related target sequences incorporated into tags 1, 2 and 3, respectively, as described in Figure 42. Tubes containing FPs in the presence of oligos comprising target sequence are indicated by asterisk.

[0043] The protocol for panel A, B, C was according to the protocol for FIG 39, and samples are described below:

Panel	FP	oligo comprising target sequence
A	FP1	TO-M1 (tube 2)
B	FP2	TO-M2 (tube 3)
C	FP3	TO-M3 (tube 4)

[0044] Figure 41 shows the sequence and predicted native conformation of fluorescent probe FP18. The sequence comprises bases which are designed to be complementary to the target sequence and additional flanking bases. The flanking bases are underlined. Panel A shows the predicted structure of the sequence using DNA folding programs according to *Nucleic Acids Res.* 31: 3429-3431 (2003). Panel B shows predicted self dimerization of the FP2 sequence according to the oligoanalyzer 3.0 software available at http://biotools.idtdna.com/analyzer/oligocalc.asp. In both Panels A and B, the

flanking bases are shaded in grey, white and black ovals indicate fluorophore and quencher moieties, respectively.

[0045] Figures 42A, B and C show the three tag sequences recognized by fluorescent probes. In Figures 42A, B and C, the target sequences are indicated in bold and they are also shown in Figure 42D. The first sequence (tag1, 42A) is the same as the sequence indicated in Figure 35B. The next two sequences (tag2, 42B and tag3, 42C) are altered versions of tag1. The differences in sequence of target2 and target3 as compared to target1 is underlined in Panel D. Additional sequence changes were made in the remaining tag sequences to compensate for the changes made in the portions shown.

[0046] Figure 43 shows the design of tag2 sequence from tag1 sequence. A-F indicate the sequential base changes made during the design.

[0047] Figure 44 shows the design of tag3 sequence from tag1 sequence. A-F indicate the sequential base changes made during the design.

[0048] Figure 45 shows that cells isolated according to the methods of this invention are viable. Panel A shows a cell isolated using FACS after the cell was transfected with three DNA constructs each encoding an RNA of interest tagged with tag1, 2 and 3, respectively. The cells were drug selected and exposed to FP1, 2 and 3 and isolated. The fluorescence intensity of the cells for each of the three probes was above background intensities compared to control cells not transfected with any of the DNA constructs. The cells were individually plated in a well of a 96-well plate directly by the FACS and one was imaged right after its isolation. Panel B shows the same cell one hour later, after it attached to the surface of the well. Panel C shows the same cell the following day, after it had undergone cell division.

[0049] The three panels each show that cells isolated according to the methods remain viable despite the previously unknown effects of the reagents used to expose the cells to the probes. These effects include compromising the plasma membrane of the cell and possibily further subjecting the cells to high pressures during FACS. Panel A shows that the cell membrane is not found to be compromised, and Panels B and C further demonstrate that the cell is viable since

it can attach to the surface of the culture dish and divide, both of which are properties of viable cells.

Detailed Description of the Invention

[0050] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. All publications and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

[0051] The term "adjacent" as used in the context of probes refers to a condition of proximity to allow an interacting pair to functionally interact with each other. For example, the condition of proximity allows a fluorophore to be quenched or partially quenched by a quencher moiety or a protease inhibitor to inhibit or partially inhibit a protease. The distance required for currently known fluorophore and quencher to interact is about 20-100 Å.

[0052] The term "basepair" refers to Watson-Crick basepairs.

[0053] The term "bulge region" refers to a single-standed region of one nucleotide or modified nucleotide that is not basepaired. The bulged nucleotide can be between mutually complementary regions (for example, Figure 4A) or may reverse the direction of a nucleic acid backbone.

[0054] The term "dumbbell structure" refers to a strand of nucleic acid or modified nucleic acid having the conformation of two stem-loop structures linked via the end of an arm from each of the stem regions (for example, Figure 7). The linkage may be a non-complementary region, or a phosphodiester linkage with or without modification.

[0055] The term "interacting pair" refers to two chemical groups that functionally interact when adjacent to each other, and when not adjacent to each other, produce a detectable signal compared to the absence of signal or background signal produced by the interacting chemical groups, or produce a different signal than the signal produced by the interacting chemical groups. An interacting pair includes but is not limited to a fluorophore and a quencher, a chemiluminescent label and a quencher or adduct, a dye dimer and FRET donor and acceptor, or a

combination thereof. A signaling probe can comprise more than one interacting pair. For example, a wavelength-shifting signaling probe has a first fluorophore and a second fluorophore that both interact with the quencher, and the two fluorophores are FRET donor and acceptor pairs.

[0056] The term "loop region" refers to a single-standed region of more than one nucleotide or modified nucleotide that is not base-paired (for example, Figure 4B). The loop can be the region in a stem-loop structure that reverses the direction of the nucleic acid backbone (for example, Figure 23A). The loop can also be between a mutually complementary region (Figure 8A)

[0057] The term "signaling probe" refers to a probe comprising a sequence complementary to a target nucleic acid sequence and at least a mutually complementary region, and further comprising at least an interacting pair. When the signaling probe is not bound to its target sequence, the moieties of the interacting pair are adjacent to each other. When the signaling probe is bound to the target sequence, the moieties of the interacting pair are no longer adjacent to each other and a detectable signal or a different signal than the signal produced by the probe in its unbound state is produced. In one embodiment, the signaling probe is a fluorogenic or fluorescent probe that comprises a fluorophore and a quencher moiety, and a change in fluorescence is produced upon hybridization to the target sequence.

[0058] The term "protease probe" refers to a probe comprising a sequence complementary to a target sequence and at least a mutually complementary region, and further comprising at least a proteolytic enzyme and at least an inhibitor of the proteolytic enzyme or another molecule capable of reversibly inactivating the enzyme. When the probe is not hybridized to a target sequence, the proximity of the proteolytic enzyme and the inhibitor of the proteolytic enzyme allows them to interact, inhibiting proteolytic activity. Upon hybridization of the probe to the target sequence, the proteolytic enzyme and its inhibitor are separated, activating the proteolytic enzyme. The proteolytic enzyme and inhibitor can be covalently or non-covalently attached to the probe.

[0059] The term "mismatch region" referes to a double-stranded region in a nucleic acid molecule or modified nucleic acid molecule, wherein the bases or

modified bases do not form Watson-Crick base-pairing (for example, Figure 4B and C). The mismatch region is between two base-paired regions. The double-stranded region can be non-hydrogen bonded, or hydrogen bonded to form Hoogsteen basepairs, etc, or both.

[0060] The term "mutually complementary region" refers to a region in a nucleic acid molecule or modified nucleic acid molecule that is Watson-Crick base paired.

[0061] The term "non-complementary region" refers to a region in a nucleic acid molecule or modified nucleic acid molecule that is not Watson-Crick base paired. For example, the non-complementary region can be designed to have bulged nucleotides, a single-stranded loop, overhang nucleotides at the 5' or 3' ends, or mismatch regions.

[0062] The term "stem region" refers to a region in a nucleic acid molecule or modified nucleic acid molecule that has at least two Watson-Crick basepairs. For example, the stem region can be designed to have more than one mutually complementary region linked by non-complementary regions, or form a continuous mutually complementary region.

[0063] The term "stem-loop structure" refers to a nucleic acid molecule or modified nucleic acid molecule with a single-stranded loop sequence flanked by a pair of 5' and 3' oligonucleotide or modified oligonucleotide arms (for example, Figure 4). The 5' and 3' arms form the stem region.

[0064] The term "three-arm junction structure" refers to a strand of nucleic acid or modified nucleic acid that has a conformation of a stem region, a first stem-loop region, and a second stem-loop region linked together via arms of the stem regions (for example, Figure 6). The first stem-loop region is 5' to the second stem-loop region. The three regions can be connected via a non-complementary region, a phosphodiester linkage, or a modified phosphodiester linkage, or a combination thereof.

Signaling Probe

Interacting Pair

[0065] The signaling probe may have more than one interacting pair, or have different interacting pairs. In one embodiment, the signaling probe is a fluorogenic probe. In one embodiment, the fluorogenic probe does not emit or emits a

background level of fluorescence in its unhybridized state, but fluoresces upon or fluoresces above the background level upon binding to its target. Multiple fluorophores can be used to increase signal or provide fluorescence at different color ranges. Multiple quenchers can be used to decrease or eliminate signal in the absence of target sequence. Examples of quenchers include but are not limited to DABCYL, EDAC, Cesium, p-xylene-bis-pyridinium bromide, Thallium and Gold nanoparticles. Examples of fluorophores include but are not limited to sulforhodamine 101, acridine, 5-(2'-aminoethyl) aminoaphthaline-1-sulfonic acid (EDANS), Texas Red, Eosine, and Bodipy and Alexa Fluor 350, Alexa Fluor 405, Alexa Fluor 430, Alexa Fluor 488, Alexa Fluor 500, Alexa Fluor 514, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 555, Alexa Fluor 568, Alexa Fluor 594, Alexa Fluor 610, Alexa Fluor 633, Alexa Fluor 635, Alexa Fluor 647, Alexa Fluor 660, Alexa Fluor 680, Alexa Fluor 700, Alexa Fluor 750, Allophycocyanin, Aminocoumarin, Bodipy-FL, Cy2, Cy3, Cy3.5, Cy5, Cy5.5, carboxyfluorescein (FAM), Cascade Blue, APC-Cy5, APC-Cy5.5, APC-Cy7, Coumarin, ECD (Red613), Fluorescein (FITC), Hexachlorfluoroscein (HEX), Hydroxycoumarin, Lissamine Rhodamine B, Lucifer yellow, Methoxycoumarin, Oregon Green 488, Oregon Green 514, Pacific Blue, PE-Cy7 conjugates , PerC, PerCP-Cy5.5, R-Phycoerythrin (PE), Rhodamine, Rhodamine Green, Rodamine Red-X, Tetratchlorofluoroscein (TET), TRITC, Tetramethylrhodamine, Texas Red-X, TRITC, XRITC. See, for example, Tyagi et al. *Nature Biotechnology* 16:49-53, (1998) and Dubertret et al., *Nature Biotechnology*, 19:365-370 (2001), incorporated herein by reference.

[0066] The invention also provides signaling probes that are wavelength-shifting. In one embodiment, one terminus of the probe has at least a harvester fluorophore and an emitter fluorophore, an adjacent terminus of the probe has at least a quencher moiety. See, for example, Tyagi et al., *Nature Biotechnology*, 18, 1191-1196 (2000), incorporated herein by reference. In one embodiment, the harvester fluorophore and the emitter fluorophore are at the same terminus, wherein the emitter fluorophore is at the distal end, and a quencher moiety is at an opposite terminus to the harvester fluorophore. The emitter fluorophore may be separated from the harvester fluorophore by a spacer arm of a few nucleotides. The harvester

fluorophore absorbs strongly in the wavelength range of the monochromatic light source. In the absence of target sequence, both fluorophores are quenched. In the presence of targets, the probe fluoresces in the emission range of the emitter fluorophore. The shift in emission spectrum is due to the transfer of absorbed energy from the harvester fluorophore to the emitter fluorophore by fluorescence resonance energy transfer. These types of signaling probes may provide a stronger signal than signaling probes containing a fluorophore that cannot efficiently absorb energy from the monochromatic light sources. In one embodiment, the harvester fluorophore is fluorescein and the emitter fluorophore is 6-carboxyrhodamine 6G, tetramethylrhodamine or Texas red.

[0067] In another embodiment, one terminus of the probe has at least a fluorophore F1, and another adjacent terminus has at least another fluorophore F2. The two fluorophores are chosen so that fluorescence resonance energy transfer (FRET) will occur when they are in close proximity. When the probe is not bound to its target sequence, upon excitation at the absorption band of F1, the fluorescence of F1 is quenched by F2, and the fluorescence of F2 is observed. When the probe is bound to its target sequence, FRET is reduced or eliminated and the fluorescence of F1 will rise while that of F2 will diminish or disappear. This difference in fluorescence intensities can be monitored and a ratio between the fluorescence of F1 and F2 can be calculated. As residual fluorescence is sometimes observed in fluorophore-quencher systems, this system may be more advantageous in the quantitative detection of target sequence. See, Zhang et al., *Angrew. Chem. Int. Ed.*, 40, 2, pp. 402-405 (2001), incorporated herein by reference. Examples of FRET donor-acceptor pairs include but are not limited to the coumarin group and 6-carboxyfluorescein group, respectively.

[0068] In one embodiment, the signaling probe comprises a luminescent label and adduct pair. The interaction of the adduct with the luminescent label diminishes signal produced from the label. See Becker and Nelson, U.S. Patent 5,731,148, incorporated herein by reference.

[0069] In another embodiment, the signaling probe comprises at least a dye dimer. When the probe is bound to the target sequence, the signal from the dyes are different from the signal of the dye in dimer conformation.

Conformation of Signaling Probes and Protease Probes

Double-stranded Structure

[0070] The present invention provides signaling or protease probes comprising at least two separate strands of nucleic acid that are designed to anneal to each other or form at least a mutually complementary region. At least one terminus of one strand is adjacent to a terminus of the other strand (Figure 1). The nucleic acid may be DNA, RNA or modified DNA or RNA. The two strands may be identical strands that form a self-dimer (Figure 8B). The strands may also not be identical in sequence.

[0071] The two separate strands may be designed to be fully complementary or comprise complementary regions and non-complementary regions. In one embodiment, the two separate strands are designed to be fully complementary to each other. In one embodiment, the two strands form a mutually complementary region of 4 to 9 or 5 to 6 continuous basepairs at each end (Figure 8, 9, 15, 22, 24 or 41). The strands may contain 5-7, 8-10, 11-15, 16-22, or more than 30 nucleotides or modified nucleotides. The two strands may have the same or a different number of nucleotides (Figure 2). For example, one strand may be longer than the other (Figure 2C). In one embodiment, the 5' end of one strand is offset from the other strand, or the 3' end of that strand is offset from the other strand, or both, wherein the offset is up to 10 nucelotides or modified nucleotides.

[0072] The region that hybridizes to the target sequence may be in the complementary regions, non-complementary regions of one or both strands or a combination thereof. More than one target nucleic acid sequence may be targeted by the same signaling probe. The one or more targets may be on the same or different sequences, and they may be exactly complementary to the portion of the probe designed to bind target or at least complementary enough. In one embodiment, the two strands form a mutually complementary region at each end and the target complement sequence resides in the regions other than the mutually complementary regions at the ends (Figure 8B)

[0073] In one embodiment, the signaling probe with at least two separate strands is a fluorogenic probe. In one embodiment, one strand has at least a quencher moiety on one terminus, and a fluorophore on an adjacent terminus of the other

strand (Figure 1). In one embodiment, each of the 5' and 3' terminus of one strand has the same or a different fluorophore, and each of the 5' and 3' terminus of the other strand has the same or a different quencher moiety (Figure 1B and 2A). In one embodiment, the 5' terminus of one strand has a fluorophore and the 3' terminus has a quencher moiety, and the 3' terminus of the other strand has the same or a different quencher moiety and the 5' terminus has the same or a different fluorophore (Figure 1C and 2B).

[0074] For the protease probe, in one embodiment, one strand has at least a proteolytic enzyme on one terminus, and an inhibitor of the proteolytic enzyme on an adjacent terminus of the other strand. In one embodiment, each of the 5' and 3' terminus of one strand has a proteolytic enzyme, and each of the 5' and 3' terminus of the other strand has an inhibitor of the proteolytic enzyme. In one embodiment, the 5' terminus of one strand has a proteolytic enzyme and the 3' terminus has an inhibitor of the proteolytic enzyme, and the 3' terminus of the other strand has an inhibitor of the proteolytic enzyme and the 5' terminus has a proteolytic enzyme.

Stem-loop Structure

[0075] In another embodiment, the signaling or protease probe is a strand of nucleic acid or modified nucleic acid that comprises at least a mutually complementary region and at least a non-complementary region. In one embodiment, the probe forms a stem-loop structure. The stem region can be mutually complementary, or comprise mutually complementary regions and non-complementary regions (Figure 4). For example, the stem region can have bulged nucleotides that are not base-paired (Figure 4). The stem region can also contain overhang nucleotides at the 5' or 3' ends that are not base-paired (Figures 3B and 3C).

[0076] When the stem region is fully complementary, the stem region can include 3-4, 5-6, 7-8, 9-10 base-pairs (Figures 3 and 5). The loop region can contain 10-16, 17-26, 27-36, 37-45 nucleotides. In one embodiment, the stem region forms 4 or 5 continuous basepairs (Figure 23).

[0077] In one embodiment, the stem-loop structure comprises at least an interactive pair comprising two chemical groups, and one chemical group is at each

terminus of the strand. In one embodiment, the signaling probe has at least a fluorophore and a quencher moiety at each terminus of the strand (Figures 3, 4 and 5). The protease probe has at least a proteolytic enzyme and an inhibitor of the proteolytic enzyme at each terminus of the strand.

[0078] In one embodiment, the stem region comprises two mutually complementary regions connected via a non-complementary region, the mutually complementary region adjacent to the interactive pair forms 5 to 9 basepairs, and the mutually complementary region adjacent to the loop region forms 4 to 5 basepairs (Figure 8, 15, 21, 22, 24 or 41). In one embodiment, the non-complementary region is a single-stranded loop region (Figure 8), a mismatch region (Figure 15) or both. In another embodiment, the stem region comprises three mutually complementary regions connected via two non-complementary regions, the first mutually complementary region adjacent to the interactive pair forms 4 to 5 basepairs, the second mutually complementary region forms 2 to 3 basepairs, and the third mutually complementary region adjacent to the loop region forms 2 to 3 basepairs.

[0079] In the stem-loop structure, the region that is complementary to the target sequence may be in one or more stem regions or loop regions, or both. The region in the stem that hybridizes to the target may be in the mutually complementary regions, non-complementary regions or both. In one embodiment, the target complement sequence is in the single-stranded loop region. In one embodiment, the regions other than the stem region adjacent to the interactive pair is the target complement sequence (Figure 8). More than one target nucleic acid sequence may be targeted by the same probe. The one or more targets may be on the same or different sequences, and they may be exactly complementary to the portion of the probe designed to bind target or at least complementary enough.

[0080] The increase in stem length may increase the stability of the signaling probes in their closed conformation, and thus, may increase the signal to noise ratio of detectable signal. Exposure of these signaling probes to cells can be carried out at slightly elevated temperatures which are still safe for the cell followed by a return to normal temperatures. At the higher temperatures, the signaling probes would open and bind to their target if present. Once cooled, the signaling probes

not bound to target would revert to their closed states, which is assisted by the increased stability of the stem. Similarly, other forces may be used to achieve the same outcome, for instance DMSO which is thought to relax base-pairing.

Three-arm Junction Structure

[0081] In another embodiment, the signaling or protease probe is a strand of nucleic acid that forms a three-arm junction structure (Figures 6A and 6B). In this structure, a stem region and two stem-loop regions are connected to form a three-way junction. The stem regions can contain 2-5, 7-9, 10-12 base pairs. The loop of the stem-loop regions can contain 3-7, 8-10, 11-13 nucleotides or modified nucleotides.

[0082] In one embodiment, the three-arm junction structure comprises at least an interactive pair comprising two chemical groups, and one chemical group is at each terminus of the strand. In one embodiment, the probe has at least a fluorophore and a quencher moiety at each terminus of the strand. The protease probe has at least a proteolytic enzyme and an inhibitor of the proteolytic enzyme at each terminus of the strand.

[0083] In one embodiment, the stem region adjacent to the interactive pair forms 3 to 4, or 3 to 6 continuous basepairs, the stem region of the first stem-loop structure forms 4 to 5 continuous basepairs, and the stem region of the second stem-loop structure forms 2 to 3 continuous basepairs (Figure 9). In one embodiment, the three regions are connected by a phosphodiester linkage or modified phosphodiester linkage via the arms of the stem regions. In one embodiment, the three regions are connected by 1 or 2 nucleotides or modified nucleotides via the arms of the stem regions.

[0084] The region in the stem that hybridizes to the target may be in the mutually complementary regions, non-complementary regions or both. In one embodiment, the target complement sequence is in the single-stranded loop region. In one embodiment, the regions other than the stem region adjacent to the interactive pair is the target complement sequence (Figure 9). More than one target nucleic acid sequence may be targeted by the same probes. The one or more targets may be on

the same or different sequences, and they may be exactly complementary to the portion of the probe designed to bind target or at least complementary enough.

Dumbbell structure

[0085] In another embodiment, the signaling or protease probe is a strand of nucleic acid that forms a dumbbell-shaped structure (Figure 7 or 11). The structure is two stem-loop regions connected via one arm of the two stem regions. The stem regions can contain 3-5, 7-9, 10-12 base pairs. The loop of the stem-loop regions can contain 5-7, 8-10, 11-13 nucleotides. In one embodiment, the dumbbell structure has one stem region of 3 continuous basepairs, and one stem region of 4 continuous basepairs. In one embodiment, the two stem regions are connected by 1 or 2 nucleotides or modified nucleotides. In another embodiment, the two stem regions by a phosphodiester linkage or modified phosphodiester linkage. In one embodiment, the stem-loop structure, dumbbell structure or three-arm junction structure has more than 30 nucleotides or modified nucleotides.

[0086] In one embodiment, the signaling probe has at least a fluorophore and a quencher moiety at each terminus of the strand. The protease probe has at least a proteolytic enzyme and an inhibitor of the proteolytic enzyme at each terminus of the strand.

[0087] The region in the stem that hybridizes to the target may be in the mutually complementary regions, non-complementary regions or a combination thereof. In one embodiment, the target complement sequence is in the single-stranded loop region. In one embodiment, the target complement sequence is the region other than the two stem regions. More than one target nucleic acid sequence may be targeted by the same probe. The one or more targets may be on the same or different sequences, and they may be exactly complementary to the portion of the probe designed to bind target or at least complementary enough.

[0088] DNA or RNA folding programs are available in the art to predict the conformation of a given nucleic acid or modified nucleic acid. Such folding programs include but are not limited to the programs described in *Nucleic Acids Res.* 31: 3429-3431 (2003) and the oligoanalyzer 3.0 software available at http://biotools.idtdna.com/analyzer/oligocalc.asp; hereby incorporated by

reference. Such folding programs often predict a number of energetically more favorable structures. In other embodiments, the invention encompasses the energetically more favorable structures of probes FP1-18 (Figures 8-24 and 41) that are predicted by folding programs. If the energy of the conformation is measured by free energy, the lower free energy value (negative) indicates that the conformation is more energetically favorable.

Chemical Modification of Signaling and Protease Probes

[0089] The present invention also provides signaling or protease probes which are chemically modified. One or more of the sugar-phosphodiester type backbone, 2'OH, base can be modified. The substitution of the phosphodiester linkage includes but is not limited to --OP(OH)(O)O--, --OP(O^-M^+)(O)O--, --OP(SH)(O)O--, --OP(S^-M^+)(O)O--, --NHP($O)_2$O--, --$OC(O)_2$O--, --$OCH_2C(O)_2$NH--, --$OCH_2C(O)_2$O--, --OP(CH_3)(O)O--, --OP($CH_2C_6H_5$)(O)O--, --P(S)(O)O-- and --$OC(O)_2$NH--. M^+ is an inorganic or organic cation. The backbone can also be peptide nucleic acid, where the deoxyribose phosphate backbone is replaced by a pseudo peptide backbone. Peptide nucleic acid is described by Hyrup and Nielsen, *Bioorganic & Medicinal Chemistry* 4:5-23, 1996, and Hydig-Hielsen and Godskesen, WO 95/32305, each of which is hereby incorporated by reference herein.

[0090] The 2' position of the sugar includes but is not limited to H, OH, C_1-C_4 alkoxy, OCH_2-CH=CH_2, OCH_2-CH=CH-CH_3, OCH_2-CH=CH-$(CH_2)_n$CH3 (n=0,1 ...30), halogen (F, Cl, Br, I), C_1-C_6 alkyl and OCH_3. C_1-C_4 alkoxy and C_1-C_6 alkyl may be or may include groups which are straight-chain, branched, or cyclic.

[0091] The bases of the nucleotide can be any one of adenine, guanine, cytosine, thymine, uracil, inosine, or the forgoing with modifications. Modified bases include but are not limited to N4-methyl deoxyguanosine, deaza or aza purines and pyrimidines. Ring nitrogens such as the N1 of adenine, N7 of guanine, N3 of cytosine can be alkylated. The pyrimidine bases can be substituted at position 5 or 6, and the purine bases can be substituted at position 2, 6 or 8. See, for example, Cook, WO 93/13121; Sanger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York (1984), incorporated herein by reference.

[0092] Derivatives of the conventional nucleotide are well known in the art and include, for example, molecules having a different type of sugar. The O4' position of the sugar can be substituted with S or CH_2. For example, a nucleotide base recognition sequence can have cyclobutyl moieties connected by linking moieties, where the cyclobutyl moieties have hetereocyclic bases attached thereto. See, e.g., Cook et al., International Publication WO 94/19023 (hereby incorporated by reference herein).

Methods

[0093] The methods of this invention are based upon the ability of signaling probes to produce a detectable signal upon hybridization to target RNA sequences in living cells. In one embodiment, the method is for detecting or quantitating RNA. One method is to isolate cells or generate cell lines that express at least an RNA. A DNA construct encoding an RNA or an RNA and a tag sequence is introduced into cells. The DNA construct may be integrated at different locations in the genome of the cell. Integration at one or more specific loci may also be accomplished. Then, the transfected cells are exposed to the signaling probe, which generates a detectable signal upon binding to the target RNA or tag sequence. The cells that produce the detectable signal are isolated. Cell lines can be generated by growing the isolated cells.

[0094] Although drug selection is not a required step, it may be used to enrich the transfected cell population for stably transfected cells, provided that the transfected constructs are designed to confer drug resistance. If selection using signaling probes is performed too soon following transfection, some positive cells may only be transiently and not stably transfected. However, this can be minimized given sufficient cell passage allowing for dilution or loss of transfected plasmid from non-stably transfected cells. Some stably integrated plasmids may not generate any RNA corresponding to cloned cDNA inserts. Others may generate RNAs which may not be or may be inefficiently detected by the signaling probes.

[0095] The RNAs can have one or more of the following different roles: messenger RNAs that encode proteins, fusion proteins, peptides fused to proteins, export signals, import signals, intracellular localization signals or other signals,

which may be fused to proteins or peptides; antisense RNA, siRNA, structural RNAs; cellular RNAs including but not limited to such as ribosomal RNAs, tRNAs, hnRNA, snRNA; random RNAs, RNAs corresponding to cDNAs or ESTs; RNAs from diverse species, RNAs corresponding to oligonucleotides, RNAs corresponding to whole cell, tissue, or organism cDNA preparations; RNAs that have some binding activity to other nucleic acids, proteins, other cell components or drug molecules; RNAs that may be incorporated into various macromolecular complexes; RNAs that may affect some cellular function; or RNAs that do not have the aforementioned function or activity but which may be expressed by cells nevertheless; RNAs corresponding to viral or foreign RNAs, linker RNA, or sequence that links one or more RNAs; or RNAs that serve as tags or a combination or recombination of unmodified mutagenized, randomized, or shuffled sequences of any one or more of the above. RNAs may be under the control of constitutive or conditional promoters including but not limited to inducible, repressible, tissue-specific, or temporal promoters or a combination or recombination of unmodified or mutagenized, randomized, shuffled sequences of any one or more of the above.

[0096] In one embodiment, the signaling probes are fluorogenic probes. Fluorescence cell sorter or related technology can be used with fluorogenic probes to identify and/or separate cells exhibiting a certain level or levels of fluorescence at one or more wavelengths. Being able to detect reliably and efficiently mRNAs as well as other RNAs in living cells enables their use to identify and, if desired, separate cells based on their desired characteristics, for instance by using a Fluorescence Activated Cell Sorter (FACS). FACS technology currently allows sorting at up to 70,000 cells per second. 5,000,000 cells can be sorted in less than 2 minutes.

1. Generating Protein-Expressing Cell Lines

[0097] Some of the most tedious steps involved in generating cell lines are eliminated by the application of signaling probes as described herein. In one embodiment, following transfection with a DNA construct encoding a desired gene, one introduces into these cells fluorogenic probes designed to recognize the message of the gene of interest. This step can be performed following drug

isolation provided that the transfected DNA construct also encodes drug resistance. Those cells transcribing the gene will fluoresce. Subsequent FACS analysis results in the isolation of the fluorescent cells which may then be grown to give rise to cell lines expressing the gene of choice.

[0098] In one embodiment, the signaling probes are designed to be complementary to either the RNA encoding the protein of interest or to their 5' or 3' untranslated regions. If the signaling probe designed to recognize a messenger RNA of interest is able to detect endogenously existing target sequences, the proportion of these in comparison to the proportion of the target sequence produced by transfected cells is such that the sorter is able to discriminate the two cell types. The gene of interest may be tagged with a tag sequence and the signaling probe may be designed so that it recognizes the tag sequence. The tag sequence can either be in frame with the message of the gene or out of frame with it, depending on whether one wishes to tag the protein produced.

[0099] Additionally, the level of expression of the gene of interest in any given cell may vary. The FACS applied here can evaluate expression levels and differentially select individual cells expressing the same gene.

2. Generating Cell Lines that Down Regulate Genes

[0100] There are several studies describing the generation of cell lines which express not RNA that encodes a protein, but rather one that is the antisense of a gene or portion of a gene. Such methods aim to reduce the amount of a specific protein in a given cell. The steps described above for the generation of protein-expressing cell lines are equally applicable here and virtually identical except that here the signaling probe is designed to detect an RNA which is an antisense RNA.

[0101] Not all attempts at making stably transfected antisense-expressing cell lines result in cell lines where the expression of the targeted protein is affected sufficiently. This difficulty has made it less worthwhile to pursue the production of such cell lines. Given the ease of the procedure described here, one easily assays the effectiveness of numerous different genetic sequences for their ability to yield active antisense expressing cell lines. One can then analyze these to

determine which exhibit appropriate expression profiles where the down regulation of targeted genes can be analyzed.

[0102] RNA interference is an alternate approach which also aims to decrease transcription levels of specific genes. RNA interference may be induced transiently using chemically synthesized siRNA or DNA constructs encoding short siRNAs, or it may be stably induced if stable cells appropriately expressing the short siRNAs are generated. The methods described here can also be used to analyze or isolate cells or cell lines based on their expression of such siRNAs.

[0103] The application of the methods described here to obtain cells expressing RNAs that induce RNA interference is additionally important because it will help overcome some complexities currently encountered in such cells. For instance, although RNA interference is performed in order to reduce the transcriptional levels of targeted RNAs and may cause downstream effects on the transcriptional levels of other RNAs, it has also been shown to reduce the transcriptional levels of RNAs that are not intended targets. The identity of these unintended targets varies depending on the sequence of the RNA that is used to induce RNA interference. This is a complicating feature of using RNA interference in cells as results may be compromised by such unintended consequences. Because the methods described here enable the efficient generation of multiple stable cells each expressing for instance, a different RNA sequence used to induce RNA interference for the same gene, each of these sequences can be assayed in the cell for its effects on the transcriptional levels of other genes. Analysis of this information can be used to distinguish RNAs that have decreased transcriptional levels, of which the decrease is not due to a decreased expression level of the targeted RNA.

3. Differentiating Between Cells Based on Cell Surface-Localized Antigens

[0104] Immunologists and others have long used the FACS to sort cells. This method is based on labeling cell-surface localized proteins with differentially labeled probes, usually fluorophore-labeled antibody probes. For instance, cells positive for expression of cell surface localized proteins may be carried out. This

method is most commonly used under conditions designed to preserve the integrity of the cell and maintain its viability.

[0105] In accordance with the present invention, to detect the presence of cell surface localized protein, a signaling probe is made to target the mRNA encoding the protein of interest. The signaling probe is introduced into cells by transfection without abrogating cell viability. Then, the cell sorter is used to isolate positive-scoring cells. Additionally, if a combination of signaling probes is used, each targeted to the mRNA of one of the proteins of interest, each can be labeled differently. If cells have a greater number of targets than what can be detected in a single application of FACS, multiple rounds of sorting are performed to sort cells.

[0106] The methods described here also enable the analysis or sorting of cells for other cellular RNAs, for instance mRNAs that code for proteins that are internally localized or secreted from the cell, or RNAs that do not code for protein. As a result, one or more of the cellular RNAs described previously may be detected as targets, including RNAs encoding proteins which are inaccessible to the commonly used antibody probes or for which probes have not yet been developed.

4. Assaying Cells for the Expression of Specific RNAs and Quantifying the Level of RNA Expression in Cells

[0107] If the target RNA of a fluorogenic probe that is introduced into a cell is present, the cell will fluoresce. This information can be qualitatively assessed by use of Fluorescence Microscopy or FACS, and it is also quantifiable by either of these. For instance, instead of performing *in situ* reverse-transcription polymerase chain reaction (RT-PCR) on slices of tissues to determine a pattern of expression for a particular RNA, a signaling probe is used to carry out the same experiment. Moreover, using a combination of differently fluorescent fluorogenic probes, each targeting a specific RNA, one assays for the presence or quantity of several RNAs of interest in one step. (Localization of Antigens in Combination with Detection of RNAs in Cells and Tissues)

[0108] Using fixed cells or tissue slices, one uses immunocytochemistry to describe the localization of the protein antigens recognized, and using signaling probe targeting specific RNAs, one co-localizes in the same samples the RNAs of

interest. It has been shown that fluorogenic probes targeted to RNAs function in fixed cells.

5. Generating Cell Lines Expressing Multiple RNAs or Proteins

[0109] Using the methods of the present invention, one very quickly generates stably transfected cell lines expressing any number of RNAs or proteins, even without the need to maintain these cells in the presence of a mixture of numerous selective drugs. Following gene transfection and optionally, drug-selection, a combination of signaling probes, one to the message for each protein, is introduced into the cells. By designing the target complementary sequence of each fluorogenic probe to hybridize to the mRNA of only one of the genes or to the tag sequences with which the messages may be associated, each signaling probe is designed to recognize the mRNA encoded by only one of the genes. In one embodiment, the cells are then sorted by FACS. By selecting for one or more signals, a variety of cell lines is generated in a single application.

[0110] One may have a need to produce a cell line expressing a number of RNAs of interest that is above the number that may be identified in a single application of FACS. For instance, it would be highly informative to have a cell line in which are over-expressed all of the proteins and RNA sequences thought to be involved in the formation of a particular complex or involved in a biological pathway. For example, RNAs or proteins in the same or related biological pathway, RNAs or proteins that act upstream or downstream of each other, RNAs or proteins that have a modulating, activating or repressing function to eath other, RNAs or proteins that are dependent on each other for function or activity, RNAs or proteins that form a complex, RNAs or proteins that share homology. If the number of RNAs required is greater than can be analyzed in one application of FACS, then to achieve this, the steps described above are repeated using cells already expressing a combination of some of the RNAs as the host cells into which would be transfected additional constructs encoding additional RNAs. Multiple rounds of the methods described may be used to obtain cells expressing the number of RNAs that is required.

[0111] If multiple RNAs to be expressed are all cloned into constructs conferring upon cells resistance to the same drug, in one embodiment, FACS can be used to isolate cells expressing all of the desired RNAs. Because the sequences are stably integrated into the genome, the cells do not lose expression of any of the sequences. However, it is possible that one or more of the sequences could be lost. If this is the case, one increases the concentration of the selective drug in the media in which these cells are grown, making this possibility less likely. Alternatively, one uses constructs each of which confers resistance to a different drug, and maintains cells in a mix of appropriate drugs. Also, a subset of the constructs to be stably transfected into cells may be chosen so as to encode a resistance gene for one drug, and another subset to encode a resistance gene for another drug.

[0112] Moreover, if some cells of a cell line lose expression of an RNA of interest, then as one resort, the first experiment to isolate the cell line as described above is repeated and new cells obtained. Alternatively, the mixture of cells described are analyzed by FACS, with the aim of re-isolating cells expressing all of the desired RNAs. This is a very useful procedure as it again yields cells which give rise to a cell line with the same genetic make-up of the original cell line selected.

[0113] The approaches described above yield an unlimited supply of cells expressing any combination of proteins and RNA sequences, amenable to virtually unlimited methods of analysis. Yet it is possible that a protein that is overexpressed may be toxic to the cell, and as will be discussed later, this possibility can be readily addressed.

[0114] The ease with which it is possible to re-isolate cells expressing all of the desired RNAs from cells which no longer express all of the RNAs makes it possible to maintain cell lines in the presence of no drug or minimal concentrations of drug. The methods described here also enable the re-application of signaling probes to cells or cell lines generated previously. For example, to determine if and to what extent the cells are still positive for any one or more of the RNAs for which they were originally isolated.

6. Generating Cell Lines Dramatically Over-Expressing One or More RNAs or Proteins

[0115] For each gene that is to be highly over-expressed, for example, two or more sequences for the same gene are first cloned into DNA constructs optionally also conferring drug resistance. Each of the multiple sequences for each gene is designed to include the sequence encoding a different tag sequence. In one embodiment, following transfection of the DNA constructs into cells and subsequent drug-selection, fluorogenic probes, each of which is targeted to only one tag sequence and differentially fluorescently labeled, are introduced into the cells and the cell sorter is used to isolate cells positive for their signals. Such cells have integrated into their genomes at least one copy of each of the differentially tagged sequences, and thus the expression of the sequence of interest occurs from an increased number of copies of essentially the same sequence of interest. The sequence of interest may be integrated at different locations of the genome in the cell. This method is used in conjunction with the use of the FACS to pick out those cells scoring most intensely for the signal of each fluorophore.

7. Generating Cell Lines Expressing Multiple Antisense RNAs

[0116] Stably transfected cell lines producing multiple antisense messages are created as follows. Such antisense messages target either mRNAs or other RNAs. One selects cells which express at different levels any one of the antisense sequences transfected. Through repeated rounds of stable transfections, one readily selects cells that would give rise to stably transfected cell lines which express the antisense message of an unlimited number of RNAs.

[0117] Of course, cells expressing other RNAs other than antisense RNAs can be prepared by the methods described herein. Such RNAs include but are not limited to one or more of mRNA, rRNA, siRNA, other structural RNAs such as hnRNA, tRNA, or snRNA, RNAs that have RNA interference activity, RNAs that serve as tags, etc.

8. Generating Libraries of Cell lines

[0118] A plurality of cells are transfected with DNA constructs that form an expression library. Expression libraries of DNA sequences can include any of the

kinds that are known in the art or a mixture of these, including but not limited to cDNA or EST libraries generated from whole organisms, tissues, cells or cell lines, and synthetic libraries including but not limited to oligonucleotides or sequences coding for peptides. Similarly, libraries of DNA constructs not specifically designated as expression libraries may be used. For instance, a DNA construct library may include sequences of DNA that may have regulatory functions such as promotor, repressor, or enhancer elements, and these may be constitutive, inducible or repressible. Likewise, DNA construct libraries may comprise large segments of DNA such as entire or partial genetic loci or genomic DNA, from which transcription may occur. Any of these expression libraries of DNA construct may each be wholly or partially mutagenized, randomized, recombined, shuffled, altered or treated in any combination thereof. Additionally, any of these types of libraries may further comprise at least one tag that is expressed and that may be used as a target for the signaling probes.

[0119] Expression libraries of DNA sequences for specific classes of proteins can be made and used to generate cell line libraries. For instance, cDNAs for protein kinases can be cloned into expression constructs comprising a tag sequence. Cells stably expressing different kinases can be obtained and used to generate a cell line library limited to cell lines expressing kinases. The class of sequences can be chosen to meet further applications such as drug screening.

[0120] Libraries of cell lines can also express classes of proteins within a sequence homology or functional family, and also proteins defined by their role in a given protein pathway or complex or system. For instance, in a drug screen for compounds which may bind to various HIV proteins, a cell line library where each cell line expresses a different HIV protein can be made and used for screening drug compounds.

[0121] Libraries of cell lines may be used to assay secreted peptides or proteins having a desired activity. First, a cell line library is made using an expression library (for peptides, proteins, ESTs, cDNAs, etc.) additionally encoding an export signal translated in frame with the peptides/proteins. Expressed peptides/proteins are secreted into the growth medium. The effect of these peptides on a particular activity can then be determined given an appropriately designed assay. Tissue

culture supernatant from the cell lines can be collected and applied to test cells to assay for the activity.

[0122] In addition, a library of cell lines may be generated for instance by transfecting an expression library into cells, and optionally, first selecting cells transfected with the expression library (for instance, using a signaling probe directed against a tag that is included in the expression library) and then exposing the cells to one or more signaling probes directed against one or more RNAs of interest. This would enable one to add RNAs to cells and determine which of these RNAs results in the downstream transcriptional upregulation or downregulation of one or more RNAs of interest.

9. Generating Cell Lines Which are Functional Knock-Outs for One or More Proteins

[0123] The methods of the present invention provide the means to prepare functional knock-outs in cultured cells. One generates cell lines which are functional knock-outs of any one protein of interest by generating cells expressing from multiple loci virtually the same antisense RNA or siRNA having RNA interference activity to a unique RNA sequence. For instance, one stably transfects into cells multiple constructs each of which would encode either the antisense RNA for a particular gene or siRNA having RNA interference activity for the gene, or both. Here each antisense RNA sequence differs only in that each would be tagged with the nucleotide sequence of a unique tag sequence. One selects those cells expressing all of the differentially-tagged antisense RNAs. Similarly, the presence of each siRNA with RNA interference activity would be determined by detection of a tag with which each siRNA is associated. Because the FACS is used to quantify fluorescence as previously described, this feature enables one to select for those cells most strongly expressing any one or more of the antisense sequences. One could isolate cells exhibiting the desired expression levels of the targeted RNA, or little or no expression of it due to expression of any number or combination of expressed RNAs, which act to decrease the expression level of the targeted RNA.

[0124] Importantly, one or more of different antisense sequences and siRNAs having RNA interference activity targeting the same gene may be used in this approach. For instance, some of the antisense RNAs, the expression of which is selected for by using signaling probes and the FACS, is designed so as to target a particular region of the messenger RNA for the gene, whereas others are designed such that they target an alternate portion of the same messenger. In order to generate cell lines which are functional knock-outs of a protein of interest, one stably transfects into cells as many genetic sequences encoding similar or different antisense RNAs or siRNAs having RNA interference activity to the same gene of interest as is necessary for the production of a cell line which exhibits no detectable levels of expression of the protein of interest, or alternatively, acceptably low levels of expression.

[0125] Moreover, one generates cell lines in which multiple proteins are functionally knocked-out or have reduced expression levels by repeating the procedure described above while targeting any number of sequences to be knocked-out functionally by antisense or siRNA. For instance, to study the function of a complex of proteins, one knock-outs or reduces the expression levels of one, all, or any combination of the proteins making up the complex.

10. Generating Cell Lines Which are Functional Knock-Outs of Only Selected Alternatively Spliced Forms of One or More Genes

[0126] Differentially spliced versions of a single gene are often translated into proteins with differing functions. Using the methods of the present invention, one generates cell lines in which only selected alternatively spliced forms of one or more proteins are functional knock-outs or are reduced in expression levels. For example, by designing antisense or siRNAs that would target only those alternatively spliced versions of the messenger RNA of the gene that one would like to eliminate from the cell, one functionally knock-outs all of the alternatively spliced RNAs of the gene of interest, or sufficiently reduces their expression levels to desired levels except for those alternatively spliced messages which are of interest.

11. Generating Cell Lines Expressing One or More RNAs or Proteins While Functionally Knocked-Out for One or More Other Proteins

[0127] For instance, for a given group of proteins that is thought to interact with each other, one can study their interactions by generating stably transfected cell lines in which one or more of the proteins of interest are functionally knocked-out or have reduced expression levels by the cell's expression of antisense or RNAs siRNAs. The function of the remaining proteins of interest in the cell can then be studied, but perhaps more interestingly, such a cell could be further altered by further manipulating it such that it will now over-express one or more of the remaining proteins of interest. In addition, one can over-express or eliminate or reduce the expression of additional proteins in cells. Again, it is possible that overexpression of certain proteins or a functional knock out of certain proteins may be lethal to cells. This is a problem that will be addressed below.

[0128] Analogous methods can be used to randomly up or down-regulate genes by introduction of DNAs with direct or indirect roles in transcriptional regulation. For instance, DNA sequences including but not limited to one or a combination of a promotor, enhancer, or repressor sequences, or a sequence which has some other binding or functional activity that results in the modulation of transcriptional levels for one or more RNAs can be transfected into cells. The activity of these elements may be constitutive, inducible or repressible. The signaling probes to one or more specific RNAs can be added to identify or isolate cells where the expression levels of these RNAs have been increased or decreased. Stable integration of some DNA sequences using the methods described may randomly sufficiently turn on or shut off transcription from genetic loci. These cell line libraries can be screened for cells which are effectively overexpressing or knocked-out for specific genes. Cells with desired levels of one or more RNAs can be selected in this way. Multiple rounds of this procedure may be performed, if necessary, to isolate cells having the desired expression profiles for multiple RNAs of interest.

12. Generation of Transgenic Mice

[0129] For some purposes, the study of cells in culture is not sufficient. The methodology described above, however, also lends itself towards the manipulation

of embryonic stem cells. Embryonic stem cells may be obtained that could either express multiple RNAs or proteins or act as functional knock-outs of multiple proteins or a subset of the alternatively spliced forms of multiple proteins, etc., following the above procedures. Such embryonic stem cells are then used as the basis for the generation of transgenic animals. Using these procedures, one determines before an attempt is made at generating an animal if the experiment will result in a lethal phenotype in the cell. For instance, if a protein that is essential is functionally knocked-out in embryonic stem cells, then such cells may not survive following their isolation.

[0130] Cells isolated according to the methods described may be implanted into organisms directly, or their nuclei may be transferred into other recipient cells and these may then be implanted. One use could be to generate transgenic animals and other uses may include but are not limited to introducing cells which synthesize or secrete cellular products into the organism and cells which are engineered to carry out desired roles in the organism.

13. Generating Inducible Stably Transfected Cell Lines

[0131] The over-expression or the lack of expression of certain proteins or RNAs in cells may be lethal or damaging. Yet it may be of critical importance to study a cell over-expressing a toxic protein or RNA, or one which is a functional knock-out of a protein or RNA, without which the cell is unable to survive. To this end, one generates stably transfected cells where selected RNAs having such deleterious effects on the cell are under the control of inducible or conditional promoters. To isolate such cell lines, in one embodiment, the transfected and optionally drug-selected cells are first minimally induced to affect transcription of the inducible genes, and the cells are then subjected to FACS analysis following the transfection into them of signaling probes designed to recognize the appropriate RNAs. The cells obtained are maintained such that the toxic RNAs are induced and transcribed only when necessary.

[0132] Inducible systems may be advantageous for applications other than the expression of toxic RNAs. For instance, one induces the expression of genetic sequences stably transfected into cells at a certain point during the cell cycle of a synchronized cell line. Alternatively, if the expressed products of a set of one or

more stably transfected genetic sequences is thought to act on the expressed products of another set, then it is of interest to clone the genetic sequences of the first set under the control of one inducible promoter, and those of the second set under the control of a second inducible promoter. By varied inductions, one studies the expressed products encoded by either set of genetic sequences in either the absence or the presence of the expressed products of the other set. In general, the DNA sequences which are incorporated into cells as described in the methods above can each be placed under the control of a promotor or other regulator of transcription with desired activity. For instance inducible, tissue-specific, time-specific or temporal promotors, enhancers or repressors may be used, as well as regulatory elements that are activated or repressed due to cellular or extracellular signals, including but not limited to one or more of compounds or chemicals, other cells, proteins, peptides, hormones, signaling molecules, factors secreted from cells, whole or fractionated extracts from organisms, tissues or cells, or environmental samples. In these cases cells are first exposed to appropriate levels of the agents regulating transcription prior to their exposure to the signaling probe.

14. Detecting Genetic Recombinational Events in Living Cells and the Subsequent Isolation of Non-Recombined or Differentially Recombined Cells

[0133] Parallel to the use of signaling probes to detect cells having undergone the recombinational events involved in the creation of stable cell lines, is the use of signaling probes to detect and isolate from a mixture of living cells those cells which have undergone other specific recombinational events. The same principle can be used to assay for VDJ recombination, translocation, and viral genome integration, for instance.

[0134] In cellular recombinational events, for instance, one sequence of genomic DNA is swapped for another. If a DNA sequence encoding a region where a recombinational event occurred is transcribed into RNA, then the presence of such an event is detected by a signaling probe designed to recognize either the RNA transcribed from the unrecombined DNA sequence, or that which is transcribed from the recombined sequence. Such an assay can also be carried out using the Fluorescent Microscope. If one would like to separate cells which have

recombined from those which have not, then one subjects the cells to FACS and sorts them. In addition, FACS can be used to sort out cells based on the presence or absence in them of numerous recombinational events.

15. Sorting Cells on the Basis of Expressed RNAs

[0135] The use of signaling probes as described herein allows cells to be sorted based on their expression of RNAs for internally localized, cell-surface localized, or secreted proteins as well as for other RNAs that may be present in the cell. For instance, starting from a mixed population of cells, one isolates those cells which express internally localized proteins of interest by designing signaling probes which recognize the mRNAs which give rise to these proteins. These signaling probes are transfected into the mixture of cells and FACS can be used to sort them as appropriate. Multiple rounds of sorting may be carried out.

[0136] Additionally, a researcher may be interested, for instance, in isolating cells which express the mRNA of one or more specific protein or RNAs that are transcribed in response to a given added factor, or in determining which added factor induces or represses the transcription of one or more proteins or RNAs. The added factors which may be tested in this way may include but are not limited to one or more of proteins, peptides, hormones, signaling molecules, chemical compounds, inorganic or organic chemicals, cells, whole or fractionated extracts from or derived from organisms, tissues or cells, samples from the environment or other sources.

[0137] To isolate cells that are induced to express one or more specific RNAs in response to a cytokine, for instance, a mixture of cells is first induced by the cytokine, then transfected with signaling probes, each of which is designed to recognize the mRNA that would give rise to one of the proteins of interest. In one embodiment, the FACS is then used to isolate those cells which score positive for the mRNA of interest. In an alternative embodiment, one also assays cells infected with a virus, for instance, for their expression of a particular gene. Alternatively, a set or library of compounds such as a chemical compound library, or an expression library of RNAs can be applied to cells to determine which if any compounds or mixture of compounds, or RNAs induces, represses or modulates the transcription of one or more specific proteins or RNAs.

[0138] It is possible with the methodology described hereinabove to detect cells positive for the presence of RNAs with one or more of the following different roles: messenger RNAs that encode proteins, fusion proteins, peptides fused to proteins, export signals, import signals, intracellular localization signals or other signals, which may be fused to proteins or peptides; antisense RNA, siRNA, short RNAs which form stem-loop structures that have an activity similar to siRNA; structural RNAs, cellular RNAs including but not limited to such as ribosomal RNAs, tRNAs, hnRNA, snRNA; random RNAs, RNAs corresponding to cDNAs or ESTs; RNAs from diverse species, RNAs corresponding to oligonucleotides, RNAs corresponding to whole cell, tissue, or organism cDNA preparations; RNAs that have some binding activity to other nucleic acids, proteins, other cell components or drug molecules; RNAs that may be incorporated into various macromolecular complexes; RNAs that may affect some cellular function; or RNAs that do not have the aforementioned function or activity but which may be expressed by cells nevertheless; RNAs corresponding to viral or foreign RNAs linker RNA, or sequence that links one or more RNAs; or RNAs that serve as tags or a combination or recombination of unmodified mutagenized, randomized, or shuffled sequences of any one or more of the above. RNAs may be under the control of constitutive or conditional promoters including but not limited to inducible, repressible, tissue-specific, or temporal promoters or a combination or recombination of unmodified or mutagenized, randomized, shuffled sequences of any one or more of the above.

16. *In Vivo* Detection of Nucleic Acids and Subsequent Selection of Cells using Protease Probes

[0139] The present invention is also directed to a novel form of protease probe which, in contrast to the signaling probe, exhibits proteolytic activity upon binding to their target nucleic acids. Such proteolytic activity may be used for detection purposes, but also to degrade particular protein sequences in a cell should the mRNA encoding the protein be present in the cell. For example, a protease which specifically cleaves a viral protein may be activated when transcription of the virus is activated, such as in a latent infection.

[0140] In another aspect, the present invention is directed to a proteolytic activity-generating unitary hybridization probe, herein referred to as a protease probe. Such protease probes also comprise nucleotides or modified nucleotides complementary to a target RNA and nucleotides or modified nucleotides mutually complementary. These protease probes operate in a similar fashion to the aforementioned probes, but instead of a production of or change in signal upon interaction of the signaling probe with its target nucleic acid sequence, the protease probe becomes proteolytic in the presence of the target.

[0141] One can substitute protease probes in place of signaling probes in the above methods, yielding new possibilities. Upon transfection of protease probes into cells expressing the RNA that is recognized by a protease probe, the protease probe hybridizes to its target. This causes activation of the protease as in its hybridized state, the protease is no longer in the vicinity of its protease inhibitor. A cell in which the target of such a protease probe is present and recognized, is damaged and is thus selected against. Cells not expressing this mRNA are more or less unaffected. Conversely, the protease probe can be designed to catalyze a cleavage reaction, which stimulates or otherwise imparts a beneficiary effect on cell growth or viability, or imparts a growth advantage to cells where its target is present and recognized.

[0142] Additionally, protease probes are useful in various other applications. The activity of the protease can be readily measured, and furthermore, the active protease in the presence of a particular nucleic acid target sequence may be employed not only for detection purposes but also for therapeutic purposes, in which, for example, a cell in which the protease probe is delivered is proteolyzed and rendered nonviable if a particular gene is transcribed, for example, one related to cellular transformation, oncogenesis, dysproliferation, and the like. For example, given a mixture of cells in which some of the cells are infected by a particular virus, one introduces into the cells a protease probes that targets a specifically viral mRNA. Cells which carry such an mRNA activate the proteolytic activity of the protease probe they contain, and this destroys these cells.

[0143] Preferably, the proteolytic enzyme inhibitor is a peptide, although other molecules including metals and metal chelators are also useful, to provide

reversible inhibition of the enzyme upon interaction with the inhibitor. Examples of useful pairs of proteolytic enzymes and inhibitors of the proteolytic enzyme include but are not limited to aminopeptidase and amastatin, trypsin-like cysteine proteases and antipain, aminopeptidase and bestatin, chymotrypsin like cysteine proteases and chymostatin, aminopeptidase and diprotin A or B, carboxypeptidase A and EDTA, elastase-like serine proteases and elastinal, and thermolysin or aminopeptidase M and 1,10-phenanthroline.

[0144] In addition, probes incorporating other interacting pairs can be used where one member of the interacting pair has a desired activity and the second acts to inhibit or diminish this activity when the probes are unbound to the target. Upon binding to their targets, the activity of the probe is exhibited as the inhibitory member of the interacting pair is no longer in the vicinity of the member having the desired activity.

[0145] Based on the foregoing description, the following methods may be carried out.

[0146] A method for isolating cells expressing at least one RNA, comprising the steps of:

a) introducing into cells DNA encoding said at least one RNA;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization to said at least one RNA; and

c) isolating said cells that produce said signal.

[0147] This method may further comprise the step of growing the isolated cells to generate a cell line expressing the RNA. A plurality of cell lines may be generated if the DNA construct is integrated at different locations in the genome of the transfected cell. Unless genomic integration of a transfected construct is directed to a particular location with the genome, integration is thought to occur randomly, so each positive cell may be different from another, and there would be multiple different cell lines all positive for the RNAs for which they selected. A plurality of cell lines may also be generated if the DNA construct is introduced into a mixed population of cells, for example, immortalized, primary, stem and germ cell lines, HeLa cells, NIH3T3 cells, HEK293 cells or CHO cells. Optionally, the

DNA construct may further encode at least a drug resistance marker, and the method may further comprise the step of selecting cells resistant to at least one drug to which said marker confers resistance after step a). Isolated cells may be grown separately or pooled. Whenever cells are isolated, whether following transfection with one or more constructs or one or more expression libraries, the isolated cells may be grown separated from each other, or pooled.

[0148] The isolated cells may be further prepared to express a second RNA. In either a simultaneous or sequential fashion, additional steps include transfecting the cell line with a second DNA construct encoding a second RNA; exposing said cells to a second signaling probe which produces a detectable signal upon hybridization to said second RNA; and isolating cells that exhibit the signal of at least one or both of said RNA and second RNA. The first signaling probe may produce the same or a differerent signal from the second signaling probe, for example, they may have the same or different fluorophores. Cell lines expressing more than two RNAs may be provided by repeating the steps simultaneously or sequentially. The second DNA construct may also contain the same or different drug resistance marker. If the first and second drug resistance markers are the same, simultaneous selection may be achieved by increasing the level of the drug. A plurality of cell lines may be generated by repeating the above steps in a simultaneous or sequential fashion using DNA constructs that form an expression library, wherein at least a portion of the cells express different RNA.

[0149] A related approach is disclosed in which a tag sequence associated with the transfected gene is used as the target for the signaling probe, of which one application is to allow the selection of cells whose RNA may be difficult to identify over background, for example, if the signaling probe detects a closely related RNA species. Accordingly, a method for isolating cells expressing at least one RNA is provided comprising the steps of:

a) introducing into cells DNA encoding said at least one RNA and at least one tag sequence;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization with the tag sequence; and

c) isolating said cells that produce the signal.

[0150] This method is essentially the same as that described previously, except that the signaling probe used is designed to recognize the tag sequence rather than said RNA. A benefit of this,procedure over the previous one is that only a small number of signaling probes, corresponding to the number of different tag sequences, is needed to prepare a large number of different cell lines expressing one or more RNAs. Optionally, the DNA construct may further encode at least a drug resistance marker, and the method may further comprise the step of selecting cells resistant to at least one drug to which said marker confers resistance after step a). Isolated cells may be grown separately or pooled. Whenever cells are isolated, whether following transfection with one or more constructs or one or more expression libraries, the isolated cells may be grown separated from each other, or pooled.

[0151] Examples of tag sequences which may be used in the invention, and to which signaling probes may be prepared include but are not limited to the RNA transcript of HA (influenza hemagglutinin protein), myc, his, protein C, VSV-G, FLAG, or FLU. These and other tag sequences are known to one of skill in the art. As used herein, the tag sequence provides at least a unique nucleic acid sequence for recognition by a signaling probe. The signaling probes have been described for use in detecting a variety of RNAs. Any of these RNAs may be used as tags. The DNA portion of the construct encoding the tag sequence may be in frame or out of frame with the portion of the DNA construct encoding the at least one RNA. Thus, the tag sequence does not need to be translated for detection by the signaling probe.

[0152] A tag sequence may comprise a multiple repeated sequence which is designed to act as target sites for a signaling probe. Such a tag sequence would provide multiple signaling probe target sites. As a result, a greater number of signaling probes can bind. This would increase the total signal that can be generated from any one nucleic acid molecule that is to be detected by a signaling probe, and thus increase signal to noise ratios.

[0153] Tag sequences can either be chosen or designed to exhibit a certain amount of predicted or experimentally determined secondary structure, with the goal of optimally presenting the tag sequence for signaling probe binding. Nucleic

acid folding prediction algorithms may be used to design potential tag sequences according to their structural adaptations. See for example, *Nucleic Acids Res.* 31: 3429-3431 (2003), hereby incorporated by reference. The nucleic acid folding prediction algorithms often predict a number of energetically most favorable structures of a given sequence. For tag sequences incorporating multiple repeated sequence units, each of these units may not necessarily adopt the same structure due to potential interactions between the repeated units and/or other sequence present in the molecule incorporating the tag sequence. The structure of any given tag sequence could be influenced by its sequence context.

[0154] In one embodiment, the tag sequence is derived from reverse-vav RNA. In one embodiment, the tag sequence forms a three-arm junction structure (Figure 36). In one embodiment, the tag sequence is 10-100, 80-100, 90-100, 80-120, 100-2Kb, 2Kb-15kb nucleotides in length. In one embodiment, the target sequence is the region from all or part of the 3' side of the stem of the first stem-loop region, to the linkage between the first and second stem-loop region, to all or part of the 5' side of the stem of the second stem-loop region (Figure 36).

[0155] In one embodiment, the stem region comprises 8-9 basepairs, the first stem-loop region comprises 4-6 basepairs and the second stem-loop region comprises 13-17 basepairs (Figure 37). In one embodiment, the stem regions of the three arms further comprise non-complementary regions. In one embodiment, the stem of the stem region and first-stem-loop region further comprise one mismatch region, the second stem-loop region further comprises 2-7 mismatch or bulge regions. In one embodiment, the linkage between the stem regions has a total of 8-12 nucleotides (Figure 37). In one embodiment, the tag sequence comprises the structure or sequence according to Figure 42 A, B or C. In another embodiment, the tag sequence has the energetically more favorable structures predicted from the sequence according to Figure 42 A, B or C.

[0156] The present invention also provides a DNA construct comprising at least one DNA encoding at least one RNA of interest and a tag sequence as described above. The invention also provides vectors and cells comprising the DNA construct.

[0157] In a further embodiment, the cell line may be made to express at least a second RNA; the steps further including transfecting the cell line with a second DNA construct encoding the second RNA and a second tag sequence, and optionally, a second drug resistance marker; optionally, selecting for cells transcribing the second marker; exposing the cells to a second signaling probe that produces a detectable signal upon hybridization with the second tag sequence, and isolating the cells that exhibit the signal of both or at least one of the first and the second tag sequence. In the case of two RNAs, the portion of the DNA sequence encoding the second tag sequence may also be in frame or out of frame with the portion of the DNA sequence encoding the second RNA. The second RNA may be transfected either simultaneously or sequentially with the first. Should the method be performed simultaneously, and the same drug resistance marker used for both constructs, a higher level of drug may be used to select for cells expressing both constructs. Furthermore, more than two RNAs may be provided in the cell line by repeating the aforementioned steps.

[0158] A plurality of cell lines may be generated by repeating the above steps in a simultaneous or sequential fashion using DNA constructs that form an expression library. In one embodiment, the expression library uses a single tag sequence, and cells are exposed to the same signaling probe which is complementary to the tag sequence. In another embodiment, different expression libraries can use different tag sequences. In one embodiment, each cell line expresses one RNA. In another embodiment, each cell line expresses more than one RNA. This can be achieved by transfecting the cell with multiple DNA constructs sequentially or simultaneously. A cell stably transfected with multiple DNA constructs may be obtained by introducing a higher concentration of DNA constructs used for transfection. Alternatively, the starting cells may have already been transfected with a first DNA construct to obtain a cell with multiple DNA constructs. Also, multiple different expression libraries may be used to transfect cells. Each library may incorporate a distinct tag sequence that is detected using corresponding signaling probes. Each expression library may comprise a drug resistance gene.

[0159] Individual or multiple cell lines may be grown separately or pooled. If a pool of cell lines is producing a desired activity, it can be further fractionated until

the cell line or set of cell lines having this effect is identified. This may make it easier to maintain large numbers of cell lines without the requirements for maintaining each separately.

[0160] Yet, another method is provided for generating a cell line that overexpresses an RNA comprising the steps of:

a) introducing into cells a first DNA encoding said RNA and a first tag sequence; and at least a second DNA encoding said RNA and a second tag sequence, wherein the first and second tag sequences are different;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization with said first tag sequence, and to at least one signaling probe that produces a detectable signal upon hybridization with said second tag sequence; and

c) isolating cells that exhibit the signal of at least one of said signaling probes.

[0161] This method may further comprise the step of growing the isolated cells to generate a cell line expressing or overexpressing the RNA. A plurality of cell lines may be generated if the DNA construct is integrated at different locations in the genome of the transfected cell. Unless genomic integration of a transfected construct is directed to a particular location with the genome, integration is through to occur randomly, so each positive cell may be different from another, and there would be multiple different cell lines all positive for the RNAs for which they are selected. Optionally, the DNA construct may further encode at least a drug resistance marker, and the method may further comprise the step of selecting cells resistant to at least one drug to which said marker confers resistance after step a). Whenever cells are isolated, whether following transfection with one or more constructs or one or more expression libraries, the isolated cells may be grown separated from each other, or pooled.

[0162] In one embodiment, the cells express an antisense or siRNA or protein. In addition, cells made to express a particular protein or proteins may be used as the starting point for creating cells expressing proteins and antisense RNA molecules. Of course, the cells expressing the antisense or siRNA molecules may be used as the starting point for adding additional RNAs encoding additional

proteins, using the methods herein. Simultaneous transfection of RNAs encoding proteins and antisense or siRNAmolecules, with corresponding signaling probes and, if desired, tag sequences, may also be performed. The various combinations of the aforementioned procedures is embraced herein.

[0163] Likewise, methods are provided for isolating cells expressing at least one RNA comprising the steps of:

a) providing cells suspected of expressing said at least one RNA;

b) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization with said at least one RNA;

c) isolating said cells that produce the signal.

[0164] The method may also be used to identify cells also expressing a second RNA, using a second signaling probe which produces a detectable signal upon hybridization to its target RNA, cells having fluorescence of both or each of the first and second signaling probes are isolated. Simultaneous expression of more than two RNAs is also achievable. Whenever cells are isolated, whether following transfection with one or more constructs or one or more expression libraries, the isolated cells may be grown separated from each other, or pooled.

[0165] Another method is provided for isolating cells expressing at least one exogenous RNA and one endogenous RNA, comprising the steps of:

a) introducing into cells DNA encoding said at least one exogenous RNA, wherein said cells potentially express at least one endogenous RNA;

b) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization to said at least one exogenous RNA;

c) exposing said cells to at least a second signaling probe that produces a detectable signal upon hybridization to said at least one endogenous RNA, wherein said second signaling probe produces a different signal than that of the first signaling probe; and

d) isolating said cells that produce at least one of said signals upon hybridization of said signaling probes to their respective RNAs.

[0166] The above two methods may further comprise the step of generating a cell line or a plurality of cell lines expressing said at least one exogenous RNA and at least one endogenous RNA by growing said isolated cells.

[0167] These methods are particularly useful for an RNA that expresses a protein that is a cell surface-localized protein, intracellular protein, or secreted protein. These methods do not require the use of probes for the proteins themselves, which may be more difficult or will affect the cell such that further experiments cannot be performed. More than one RNA encoding a protein can be identified using a plurality of signaling probes, up to the number simultaneously detectable by the technology used for isolation. Optionally, the DNA construct may further encode at least a drug resistance marker, and the method may further comprise the step of selecting cells resistant to at least one drug to which said marker confers resistance after step a). Isolated cells may be grown separately or pooled. Whenever cells are isolated, whether following transfection with one or more constructs or one or more expression libraries, the isolated cells may be grown separated from each other, or pooled.

[0168] For the above methods, in one embodiment, the cells are implantable in an animal. In one embodiment, the signaling probe is a fluorogenic probe, and the cells that express said RNA fluoresce. Isolating said cells that fluoresce may be carried out using fluorescence activated cell sorter technology. In one embodiment, two signaling probes are used to target the RNA or tag sequence. The fluorophore of the first probe may be the same or different from that of the second probe. In the case of two different fluorophores, they may have similar or different emission wavelengths. At present, FACS technology can allow the detection of up to seven different fluorophores during a sorting procedure. The above methods may be repeated simultaneously to obtain the expression of up to seven different proteins. If desired, a cell line expressing seven different proteins may then be used as the starting point for the introduction of more proteins following the procedure. As FACS technology advances, it will be able to resolve a greater number of signals, and one would be able to select for cells having a greater number of RNAs in one application of FACS.

[0169] Naturally, the aforementioned procedures may be used to quantify the level of at least one RNA transcript expression in a biological sample comprising the steps of:

a) exposing the biological sample to a first signaling probe which produces a detectable signal upon hybridization with said RNA transcript;

b) quantitating the level of the signal in the biological sample; and

c) correlating the level of signal with said level of the at least one mRNA transcript.

[0170] The biological sample may be a cellular sample, a tissue sample or preparations derived thereof; these may be fixed, for example, with formaldehyde, glutaraldehyde, or any number of known cellular fixatives which do not interfere with the detection of RNA using signaling probes. Preparations of cellular samples include but are not limited to subcellular organelles or compartments, mitochondria, organelles of cells, subcellular fractions, plasma membrane of cells intracellular or extracellular material, membrane preparations, preparations of nucleic acid from any one or more of these, preparations of any virus or other virus or organism present in any one or more of these, or any combination of one or more of these.

[0171] For the embodiment of fluorogenic probes, the fluorescence may be quantitated by fluorescence microscopy or fluorescence-activated cell sorter technology. Additional RNA species may be quantitated simultaneously using a second signaling probe which fluoresces upon hybridization to a second RNA transcript. The above method may be used simultaneously with assays that utilize a fluorogenic reporter for the detection of intracellular events, states or compositions. Such fluorescent assays include but are not limited to TUNEL, Apoptosis, necrosis, Ca2+/Ion flux, pH flux, immunofluorescence, organelle labeling, cell adhesion, cell cycle, DNA content, interactions between: protein-protein, protein-DNA, protein-RNA. Reagents which may be fluorescently labeled for use in these assays include but are not limited to Proteins (labeled with fluorescent molecules or autofluorecent proteins); fluorescent metabolic indicators

(C12 resazurin, CFSE for cell divisions); fluorescent substrates or by-products; fluorescently-labeled lectins; fluorescent chemicals; caged fluorescent compounds; fluorescent nucleic acid dyes; fluorescent polymers, lipids, amino acid residues and nucleotide/side analogues.

[0172] Introduction of antisense or siRNA molecules in cells is useful for functionally eliminating or reducing the levels of one or more proteins or RNAs from the cell. Following the above methods, a method is provided for isolating cells or generating cells functionally null or reduced for expression of at least one preselected protein or RNA comprising the steps of providing in said cells a plurality of antisense or siRNA to said preselected protein or RNA, each provided in accordance with the aforementioned methods, wherein said plurality of antisense or siRNA binds essentially all or a sufficient level of mRNA transcripts of said at least one preselected protein or RNA. The preselected protein may be an alternatively spliced form of a gene product.

[0173] Following similar lines, a method is provided for generating a transgenic animal that is a functionally null-expressing mutant of at least one preselected protein or RNA, or that expresses said at least one preselected protein or RNA at reduced levels, comprising carrying out the steps described hereinabove utilizing embryonic stem cells, determining the viability of said stem cells functionally null or reduced for expressing said preselected protein or RNA, and using said viable embryonic stem cells to produce said transgenic animal.

[0174] Likewise, a method is provided for isolating cells or generating a cell line which is functionally null or reduced for expressing at least one protein or RNA and overexpresses at least one other protein or RNA, comprising carrying out the methods herein on the same cells. In similar fashion, a method is provided for generating a cell line expressing a lethal antisense or siRNA under control of a inducible promotor, or a sequence which has some other binding or functional activity that results in the modulation of transcription levels. This can be achieved by carrying out the method herein, wherein the transfection step is performed in the presence of a minimal amount of an inducer or compound.

[0175] Therefore, the present invention provides a method for isolating cells that overexpress at least a first protein and which are functionally null expressing or reduced in expression for at least a second protein, comprising the steps of:

a) introducing into cells at least a first DNA encoding at least one RNA that encodes said at least first protein, and at least a first tag sequence; and at least a second DNA encoding said at least one RNA and at least a second tag sequence, wherein said first and second tag sequences are different;

b) introducing into cells at least one DNA encoding at least one antisense RNA or siRNA that binds to or interferes with the mRNA transcript of said at least second protein;

c) exposing said cells to at least a first signaling probe that produces a detectable signal upon hybridization with said at least first tag sequence, and to at least a second signaling probe that produces a detectable signal upon hybridization with said at least second tag sequence;

d) exposing said cells to at least one signaling probe that produces a detectable signal upon hybridization to said at least one antisense RNA or siRNA; and

e) isolating cells that produce at least one of said signals upon hybridization of said signaling probes to their respective RNAs.

[0176] In yet another embodiment, the present invention provides a method of identifying a compound that modulates transcription of at least one preselected RNA, comprising the steps of:

a) adding individual or a set of compound to cells;

b) exposing said cells to at least one signaling probe which produces a detectable signal upon hybridization with said at least one preselected RNA;

c) quantitating the level of the signal in said cells;

d) identifying cells that have an increase or decrease in signal compared to the signal of cells with no compound added; and

e) identifying compounds that modulate transcription of said at least one preselected RNA.

[0177] In one embodiment, said preseleted RNA is encoded by the genome of the cell. In one embodiment, said preselected RNA is encoded by a DNA construct that is transfected into the cells prior to step a). In one embodiment, the transfected cells are exposed to a signaling probe designed to recognize said RNA and the cells express said RNA. In one embodiment, the DNA construct comprises a promotor or operator and encodes a repressor, enhancer, or a sequence that modulates transcription. In one embodiment, the preselected RNA is linked to a tag sequence, and the signaling probe produces a detectable signal upon hybridization with the tag sequence.

[0178] In another embodiment, the present invention provides a method of identifying an RNA sequence that modulates transcription of at least one preselected RNA, comprising the steps of:

a) introducing into cells at least a construct encoding a test RNA sequence that potentially modulates transcription of said at least one preselected RNA;

b) exposing said cells to at least one signaling probe which produces a detectable signal upon hybridization with said at least one preselected RNA;

c) quantitating the level of the signal in said cells;

d) selecting cells that have an increase or decrease in signal compared to the signal of cells with no test RNA sequence; and

e) identifying a test RNA sequence that modulates transcription of said at least one preselected RNA.

[0179] These cells can be isolated and grown to give rise to cell ines. They may be grown separately or pooled. The modulation of transcription can be downstream up or down regulation of the preselected RNA. In one embodiment, said preselected RNA is encoded by the genome. Said preselected RNA is encoded by a DNA construct that is transfected into the cells prior to step a). In one embodiment, the transfected cells are exposed to a signaling probe designed to recognize said RNA and the cells express said RNA. In one embodiment, the preselected RNA is linked to a tag sequence, and the signaling probe produces a detectable signal upon hybridization with the tag sequence. In one embodiment, an

expression library of RNA sequences is used to identify the RNA sequences that modulate transcription. In one embodiment, the test RNA sequence is linked to a tag sequence. Step e) is facilitated by exposing the cells following step a) to at least a signaling probe that produces a detectable signal upon hybridization to said RNA sequences or tag sequences, followed by step b). The signals produced by the signaling probes directed to the test RNA sequence or tag sequence thereof may be different from the signaling probe directed to the preselected RNA, and therefore the cells may be exposed to different signaling probes simultaneously.

[0180] A method is also provided herein for identifying genetic recombinational events in living cells comprising the steps of:

a) exposing a cell to a signaling probe that produces a detectable signal upon hybridization with an RNA sequence selected from the group consisting of that transcribed from a recombined sequence and that transcribed from the non-recombined sequence; and

b) detecting said cells expressing said RNA sequence.

[0181] The detecting and/or sorting of the cells may be performed by FACS or fluorescence microscope.

[0182] The present invention may be better understood by reference to the following non-limiting Examples, which are provided as exemplary of the invention. The following examples are presented in order to more fully illustrate the preferred embodiments of the invention. They should in no way be construed, however, as limiting the broad scope of the invention. Methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention.

Example 1

[0183] General Protocol. Starting Material: Signaling probes may be introduced into cells which are not expressing any RNAs from the DNA construct, or they may be used to detect RNA messages encoded from the DNA construct. The method of introduction of the signaling probes into either of these two types of cells is identical. The protocol below only requires that the cells to be analyzed are separable from each other and are amenable to FACS analysis.

1) As described more thoroughly in the description of the invention, signaling probes can be used in conjunction with FACS to sort out cells of a tissue based on expression or lack of expression within cells of specific RNAs. To this end, cells are first separated from each other by standard and well established methods such as homogenization and further chemical treatment. Appropriate signaling probes may then be introduced into such cells according to the protocol below.

2) Second, one may use signaling probes to select for cells expressing particular RNAs encoded by the DNA construct that have been transfected into a population of cells. To this end, one first transfects into a culture of cells a DNA construct or DNA constructs encoding the desired RNAs. Signaling probes may then be generated to recognize these RNAs, as described in more detail in the description of the invention. Transfection of the DNA construct into cells can be accomplished through a vast variety of methods using either ones own reagents or kits obtained from biotechnical firms (Qiagen, Promega, Gene Therapy Systems, Invitrogen, Stratagene, etc.), following the manufacturers' instructions. The DNA constructs are chosen such that each confers resistance to an antibiotic. Following the transfection of these DNA constructs into cells and a brief period for the recovery of the cells (usually 24 hours), the cells are subjected to the appropriate antibiotics such that only those cells to which the DNA constructs have conferred antibiotic resistance will survive. This generally takes three to four days and sometimes longer, depending both on the cell type and the antibiotic used.

[0184] The result is that a pool of cells remain and all of these would be resistant to antibiotics, but only a small fraction of which express the RNAs of interest. To select for the cells expressing the desired RNAs, the protocol below may be followed.

Example 2
Selection of Cells Using Signaling Probes

1) Transfect signaling probes into cells: signaling probes must be designed such that they will recognize the desired RNA either by hybridizing to a sequence endogenous in the RNA or by hybridizing to a tag that is

added to the native RNA sequence. The design of signaling probes is elaborated upon in the description of the invention.

[0185] Transfection may be carried out by a vast variety of methods, similar to the transfection of the DNA constructs into cells. The method employed should be chosen based on the cell type being used as some cells respond better to some transfection methods over other methods. Transfection should be performed according to the instructions of the manufacturer of the transfection reagent used and may need to be optimized.

[0186] Transfection of signaling probes into cells may be carried out either on cells in suspension or on cells growing on solid surfaces, depending on the cells and transfection reagent used.

2) Following the transfection of fluorogenic probes into cells, the cells may then be subjected to FACS analysis. FACS can be used to sort out cells positive for any one or more of the fluorogenic probes used. It can also be used to sort out cells based on the intensity of the fluorogenic probes' signal, thereby allowing the researcher to select cells which express RNAs at varying levels.

Example 3
Generation of Cell Lines Expressing One or More RNAs

[0187] Following FACS selection, the positive-scoring cells can be maintained in appropriate medium as described in more detail in the description of the invention. These cells would give rise to cell lines expressing the RNAs of interest.

[0188] Concentration of the signaling probe: The concentration of signaling probe to be used depends on several factors. For instance, one must consider the abundance within cells of the RNA to be detected and the accessibility of this RNA to the signaling probe. For instance, if the RNA to be detected is present in very low amounts or if it is found in a portion of the RNA which is not readily accessible based on the three-dimensional folding of the RNA or due to protein binding to the RNA, then more signaling probe should be used here then in cases where the RNA to be detected is in high abundance and where the site recognized by the signaling probe is fully accessible. The exact amount of signaling probe to be used will have to be determined empirically for each application.

[0189] This can be accomplished by introducing different amounts of signaling probes into different groups of cells and selecting the condition where background fluorescence is low and where signal is high (the condition where not all but some of the cells score positive for the signaling probe).

Example 4

Exposing Cells to Signaling Probes

[0190] Cells attached, partially attached to, or settled on surfaces, or in solution may be exposed to FPs using a variety of methods to introduce molecules into cells, including but not limited to methods known in the art such as microinjection, mechanical shearing forces such as vortexing or mixing, passing through needles, or cell loading techniques including scraping, permeabilization using reagents such as certain antibiotics or detergents or a combination of reagents or solvents, or through use of a variety of transfection reagents of varying chemical properties (for instance liposomal based, chemical, or protein based), or through a combination of any one or more of these methods.

[0191] A more detailed sample protocol describing the use of lipid based transfection reagents is outlined below:

1. Preparation of Cells

[0192] Cells were plated into tissue culture wells either prior to (cell plating method 1) or on the same day as their exposure to (cell plating method 2) FPs, or cells were transferred to microcentrifuge tubes on the same day as their exposure to FPs (cell plating method 3).

[0193] All three preparative methods were used successfully. Cell plating method 1 allows sufficient time for the cells to attach to the surface of the culture well, while cell plating method 2 allows cells settled on the plate or attached to varying degrees to be processed depending on how much time is allowed to pass before cells are further processed, and cell plating method 3 allows for cells to be processed directly after they have been transferred to the tubes without allowing any time for them to settle or attach to any surface, although processing may also be carried out after a given amount of time has allowed to pass.

[0194] The cells were rinsed once or more with buffer such as serum-free media or PBS, although other buffers may be used. Generally, buffers included $MgCl_2$ at varying millimolar concentrations. The rinsing step may be omitted depending on the method of exposure.

2. Preparation of FP Reagents and Exposure to Cells

[0195] FPs were prepared for addition to cell preparations using commercially available transfection reagents and following protocols as described by the manufacturers. Manufacturers' protocols instruct that multiple parameters need to be empirically determined for multiple variables including which cell types are used, at what confluency or concentration cells should be processed, which specific molecules are to be introduced, in which proportion and absolute amounts various reagents are to be combined, and for what durations various steps are to be carried out or incubated, which steps may be omitted, and other variables. In the manufacturer's protocol, while certain ranges are provided, the protocols also suggest that these may have to be exceeded or other parameters may have to be optimized for successful use of their reagent. In general, the exposure of cells to FPs was carried out using parameters for these conditions which were within the ranges suggested as preliminary ranges by the manufacturers' protocols.

[0196] In general, FPs would be added to a tube containing serum-free media and the transfection reagent would be added to a second tube also containing serum free media using volumes and concentrations suggested by the manufacturer. The contents of each tube would be mixed, combined, and incubated for a length of time all as indicated by the manufacturers' protocols. Next, the FP would be applied to the preparations of cells, and the cells would be assayed following various incubation periods, all according to manufacturers' protocol.

[0197] The protocol for exposing the FPs to cells in Figures 32 is as follows:

[0198] Cells were plated a day before exposure to FPs, at approximately $1x10^5$ cells/ml and at 0.5ml per well of a 24-well plate. Cells used were cells transfected with and drug selected for an expression construct encoding r-vav.

[0199] FP reagents were prepared by incubating 0.625 to 2.5ul of a 20uM stock of FP in 50ul to 200ul of serum-free media containing from 1 to 4 mM $MgCl_2$ in

one tube, and 0.625 to 2.5ul of TfX50 (Promega) in an equal volume of serum-free media having the same concentration of $MgCl_2$ in a second tube. The contents of each tube would be mixed and then combined and incubated for 15 to 45 minutes at room temperature.

[0200] Cells were rinsed one or more times with serum-free media supplemented to the same concentration of $MgCl_2$ used above. The preparation of FP would be applied to the cells and extra serum-free media plus $MgCl_2$ may optionally be added depending on the volume of FP preparation added such that the cells are covered.

[0201] Cells were incubated for 3 to 5 hours in a tissue culture incubator and then assayed, and optionally DMSO may be added prior to observation.

[0202] The protocol for Figure 34 was carried out essentially as described for Figure 32 except FP16 was added to cells that was either transfected or untransfected and drug selected for an expression construct encoding an RNA comprising the target sequence 6CA4.

[0203] The protocol for exposing the FPs to cells in Figure 33 (A,B,C) are as follows:

[0204] Approximately 50 to 100ul of cells plated at approximately $2.5x10^5$ cell/ml in PBS supplemented with 4mM $MgCl_2$ (PBS+4) was plated in wells of a 96-well plate, and the cells were exposed to FPs after sufficient time had passed for them to settle on the plate surface but before they had spread.

[0205] FP reagents were prepared by mixing 2.5ul of a 20uM stock of FP in 100ul of PBS+4 in one tube and 7ul of Lipofectamine (InVitrogen) was added to 100ul of PBS+4 in a second tube. The contents of each tube were mixed and the solutions were incubated for 30 minutes at room temperature before they were combined, mixed and incubated for an additional 15 minutes at room temperature. 50ul of the FP preparation was added to each well. A rinse of the cells with serum-free media or other buffer is optional. Cells were assayed following incubation in a tissue culture incubator for 2 to 3 hours and then assayed.

[0206] The protocol for exposing the FPs to cells in Figure 33 (D,E) were essentially the same as for Figure 32 except 4ul of Plus Reagent (InVitrogen) was added to the tube containing the FP.

Example 5 Analyzing and Isolating Cells via FACS

[0207] Cells were exposed to FPs as described above and then processed to detach them from surfaces if they were attached and separate them from each other (for instance by using trypsin, although other enzymatic or non-enzymatic procedures may be used) and then they would be applied to FACS.

[0208] Following exposure to FPs, the FP containing solution would be removed from the cells and trypsin would be applied directly to the cells. Rinsing of the cells prior to this step using a variety of buffers or reagents may be used, including reagents designed to remove any FP reagent or other reagent which may have associated with cell surfaces during the exposure of the cells to FPs. The cells would be resuspended in buffer (for instance media containing serum and magnesium), and the cells would be further dispersed for instance by pipeting.

[0209] The cells were then analyzed by FACS according to standard FACS methods. For analysis of the cells, by comparing the fluorescence intensities of control cells and cells potentially expressing target sequences for the presence of these sequences using FPs, one can determine the background fluorescene of cells having undergone this procedure to determine if the cells potentially expressing targeted sequences show any changed fluorescence. Cells can be isolated either individually or in batch based on this information. For instance, existing technology enables the direct isolation of desired cells into unique wells of 96-well plates, or multiple desired cells may be obtained.

[0210] By increasing or decreasing the threshold of fluorescence intensity in FACS, cells exhibity different levels of fluorescence signal may be isolated. One may wish to either set the threshold very high to have the high assurance that the cells are indeed positive or one may set a lower threshold if the purity of positive cells in the population that is obtained is not critical, for instance in the case where one would simply like to enrich the isolated cells for positive cells.

Example 6 Determining the Stability of Expression

[0211] One can monitor the expression levels of one or more genes in cells over time. For instance, given a cell line derived from cells isolated as positive for

expression of a sequence using a FP designed to recognize the sequence, one can expose the cells and control cells to the FP and determine the ratio of fluorescence intensities of the two cell types for the signal emitted by the FP. This procedure can be repeated over time and changes in the ratio will reflect changes in relative levels of expression for the sequence being detected by the FP. This procedure may be carried out using multiple FPs.

Example 7 Design of Tag Sequence

[0212] The sequence for tag1 was used to generate two additional different sequences such that each of the sequences could be recognized by a unique signaling probe, with the intention of generating sequences having the same or similar predicted structure particularly with respect to the regions of the structure most directly involved in binding the signaling probe such that the target sequence of the signaling probe would be similarly presented for binding. For each of the two different tag sequences (tag2 and 3) that were generated, the targeted sequence within the original sequence for tag1 was first changed, and compensatory changes were made to some additional bases predicted to interact with the changed sequences in effort to preserve the interactions at these same positions (Figure 43 and 44). The predicted structure of the new sequence was obtained and if it did not closely match the predicted structure of tag1, then the structure was used to predict which additional changes would be necessary. This was an iterative process performed until novel sequences having predicted structures similar to that of tag1 were obtained. Changes made to these sequences include base substitutions, deletions and additions.

Example 8 Exposing Signaling Probes at Elevated Temperatures

[0213] FP17 is predicted to form a 7 base pair mutually complementary region adjacent to the interactive pair. Elevated temperatures was required for FP17 to give a stronger fluorescence signal when it was incubated in the presence of target oligo compared to control oligo. For instance, by briefly placing them in hot water of 90 to 95°C. The tubes used here contained 16ul total consisting of 5ul of a 20uM

FP stock, 1.5ul 25mM $MgCl_2$, 8ul 20uM oligo, and 1.5ul of water, having a final magnesium concentration of approximately 2.34mM. The target oligo used was TO-FP1 and the control oligo used was TO-FP18 as described above.

Abstract

[0214] The present invention relates to methods of isolating cells or generating cell lines using signaling probes that produce a signal upon hybridization to a target sequence. Other methods that utilize the signaling probe include methods of quantifying the level of RNA expression, methods for identifying genetic recombinational events in living cells and methods of generating a transgenic animal using the isolated cells. The invention also provides protease probes. Signaling probes and protease probes that form stem-loop structures, three-arm junction structures, and dumbbell structures are provided.

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

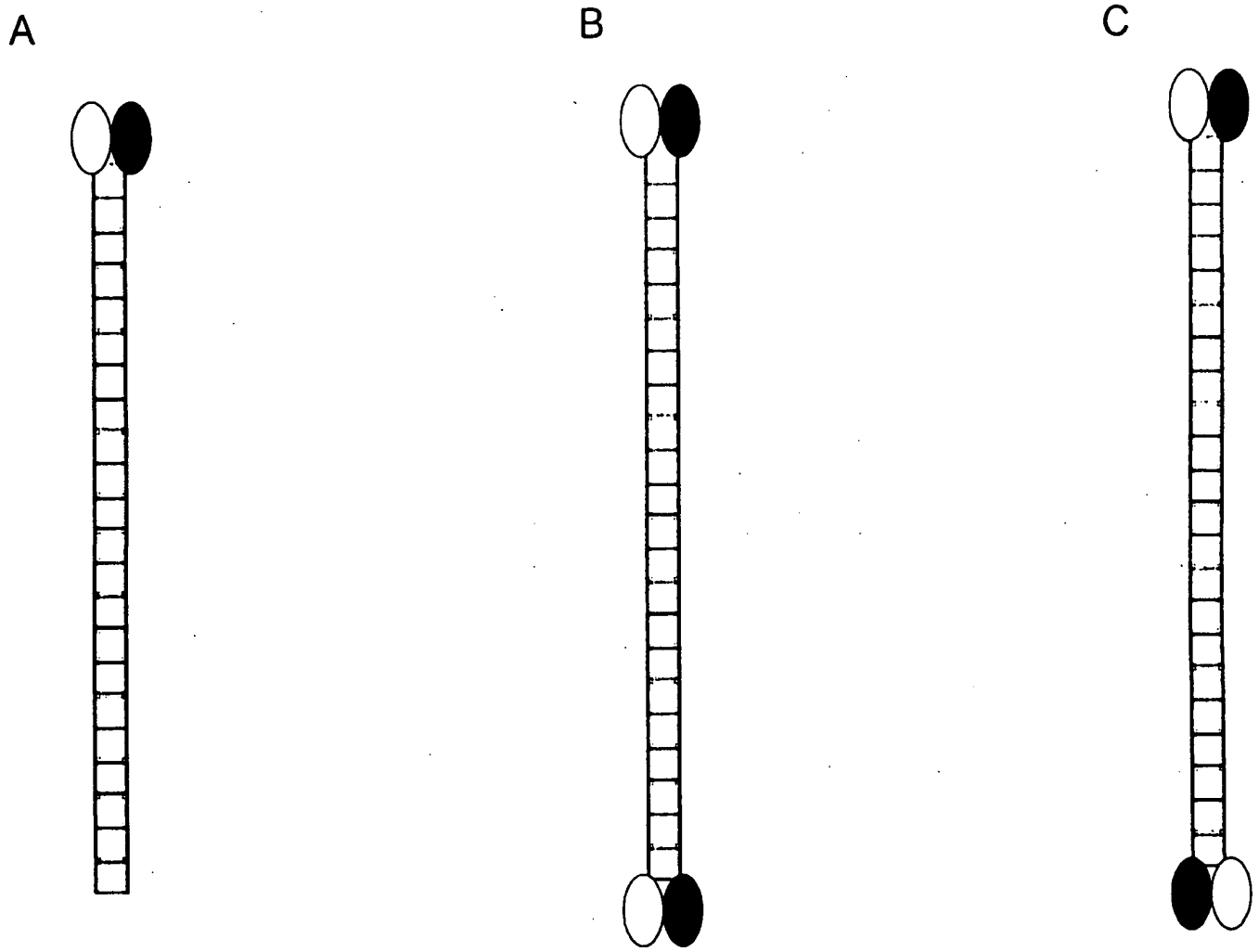


FIG. 1

Express Mail No.
EV079959879US

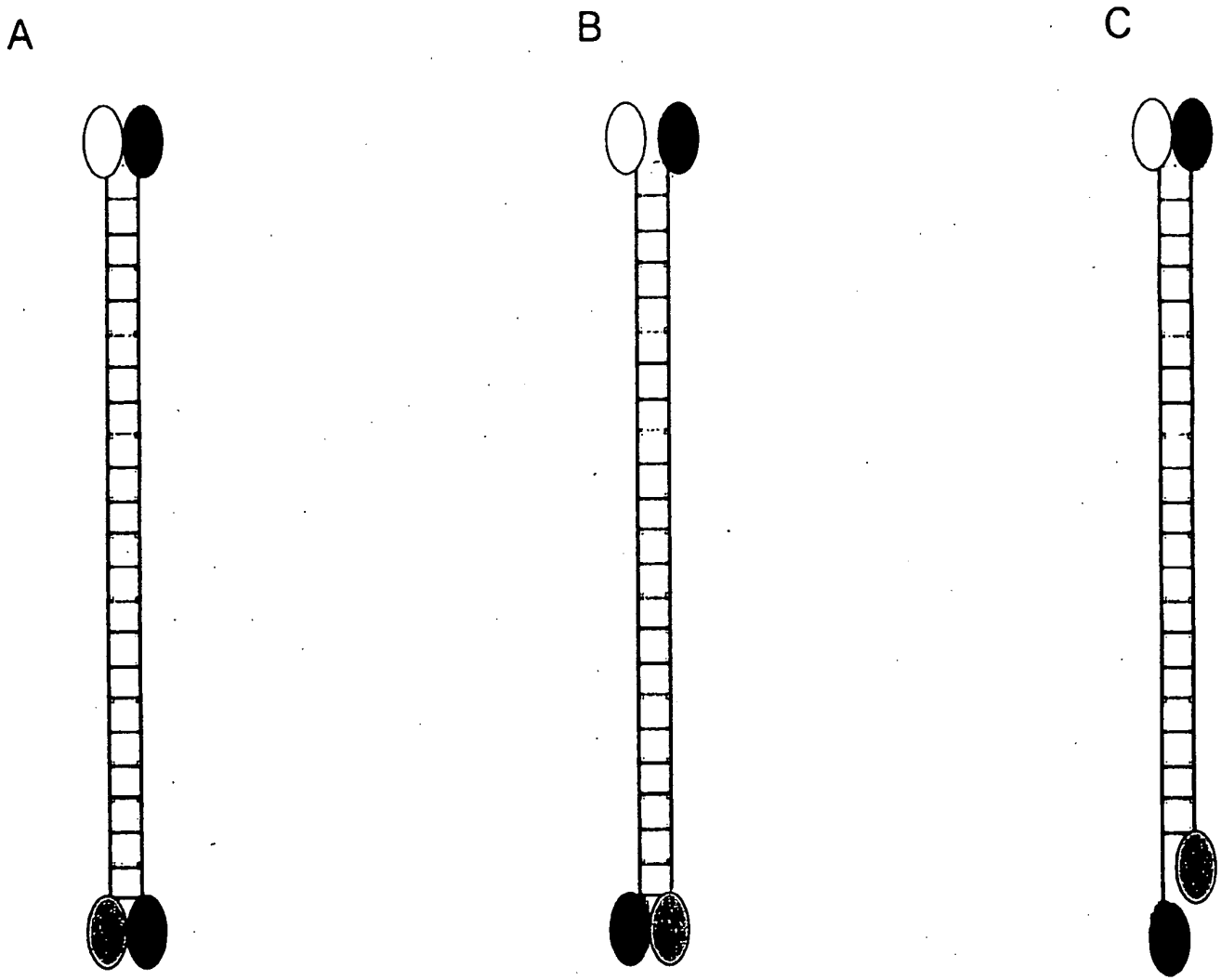


FIG. 2

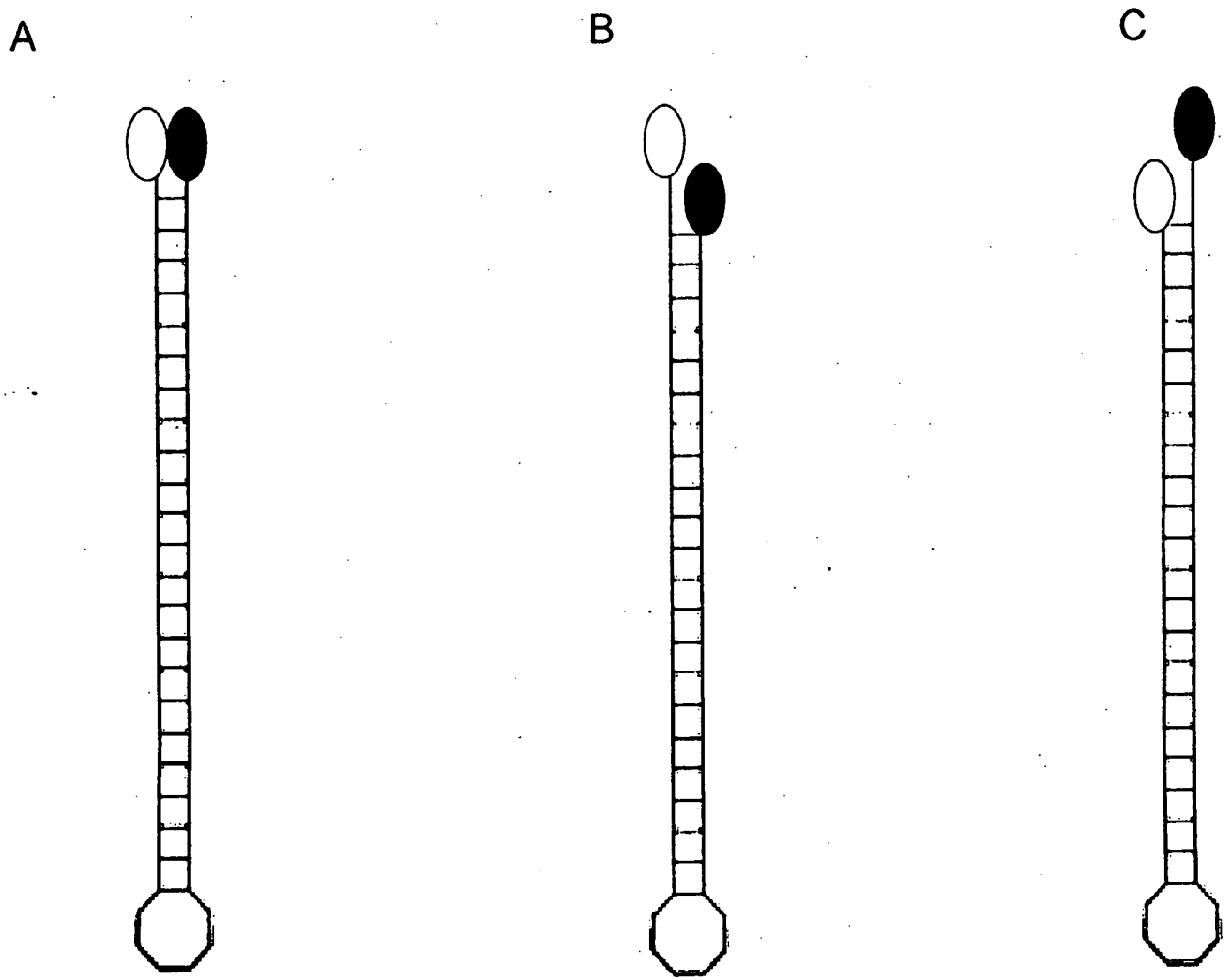


FIG. 3

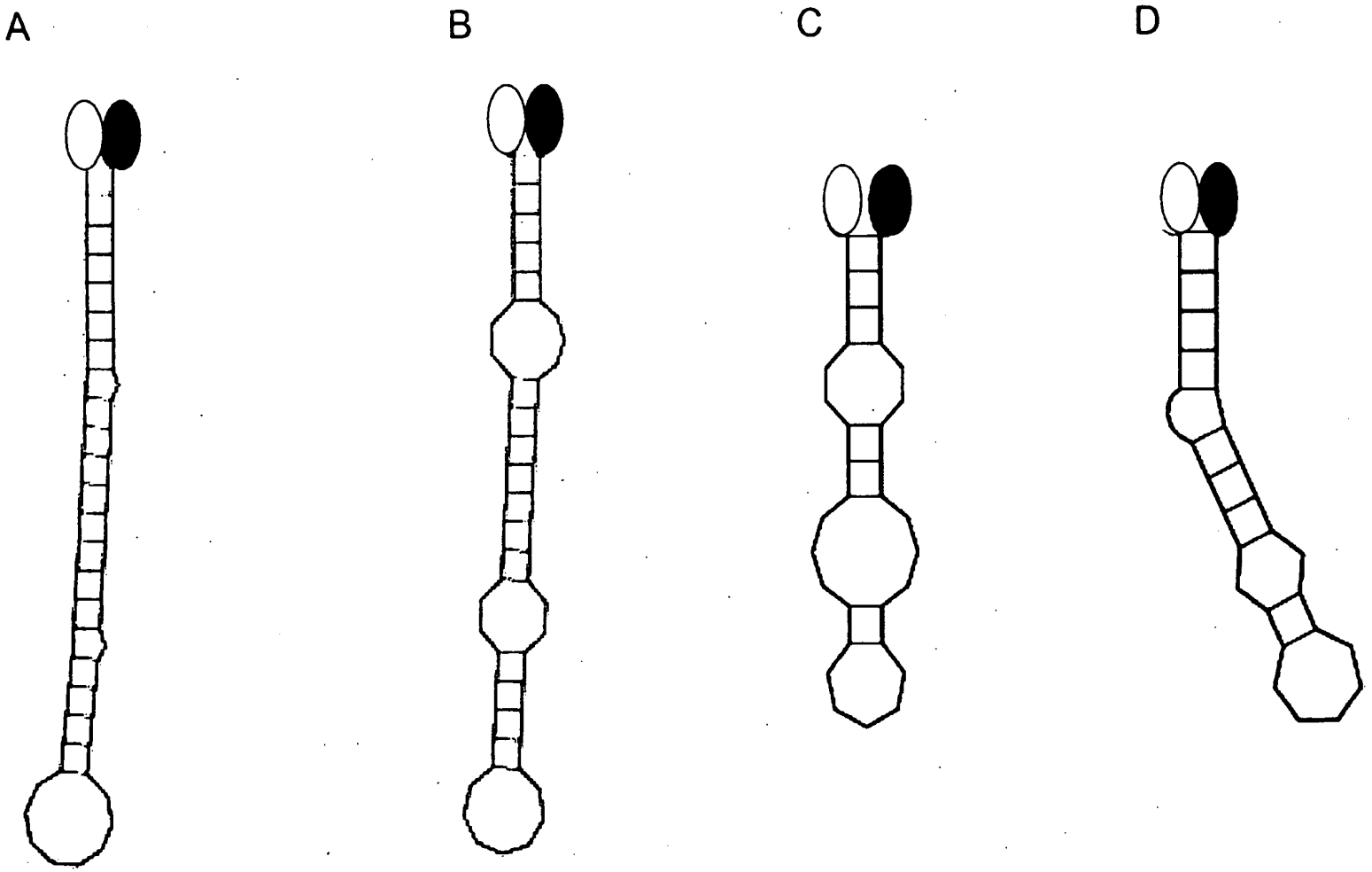


FIG. 4

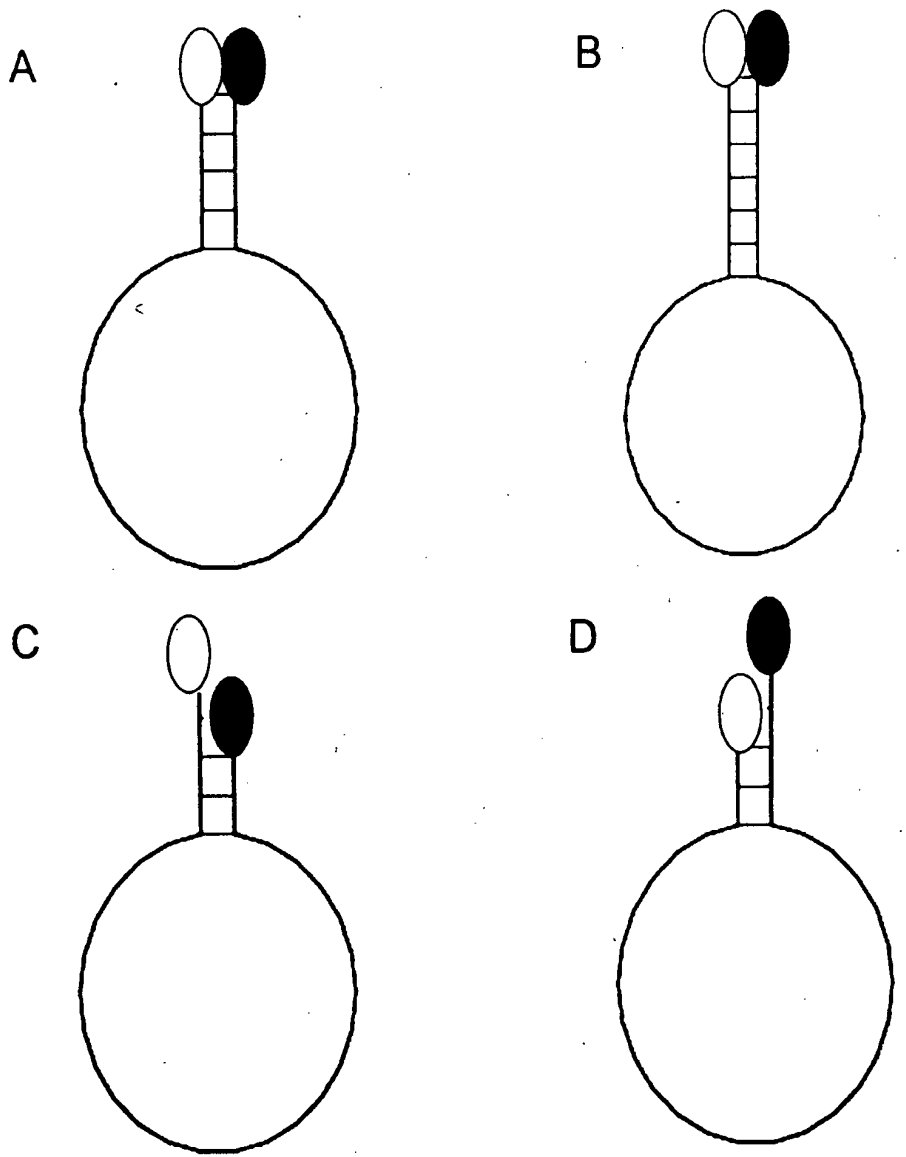


FIG. 5

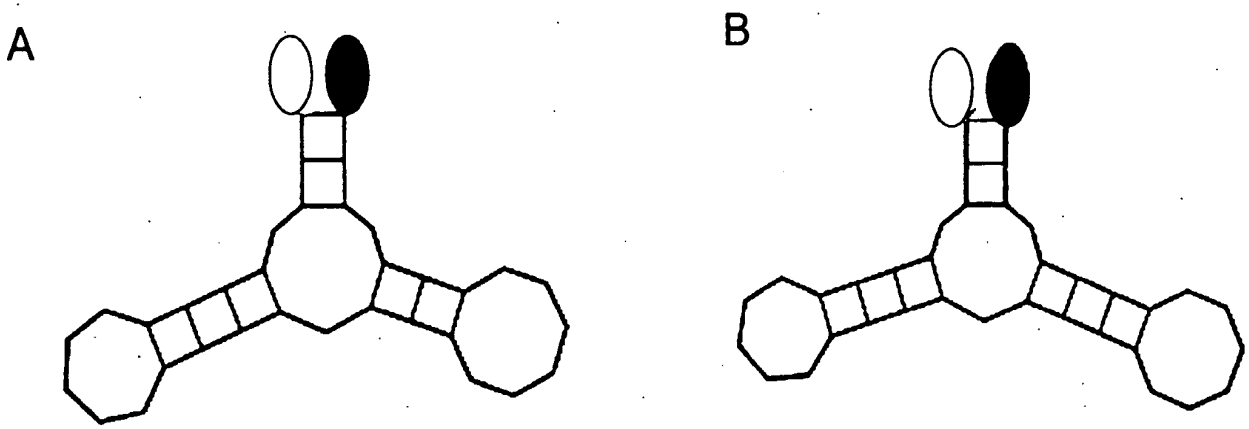
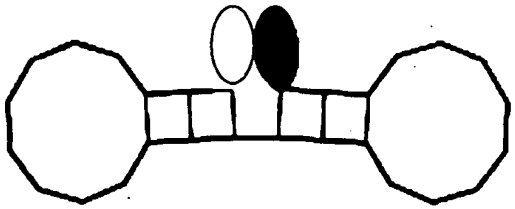


FIG. 6

A



B

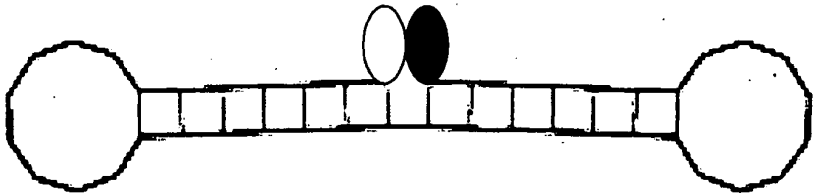
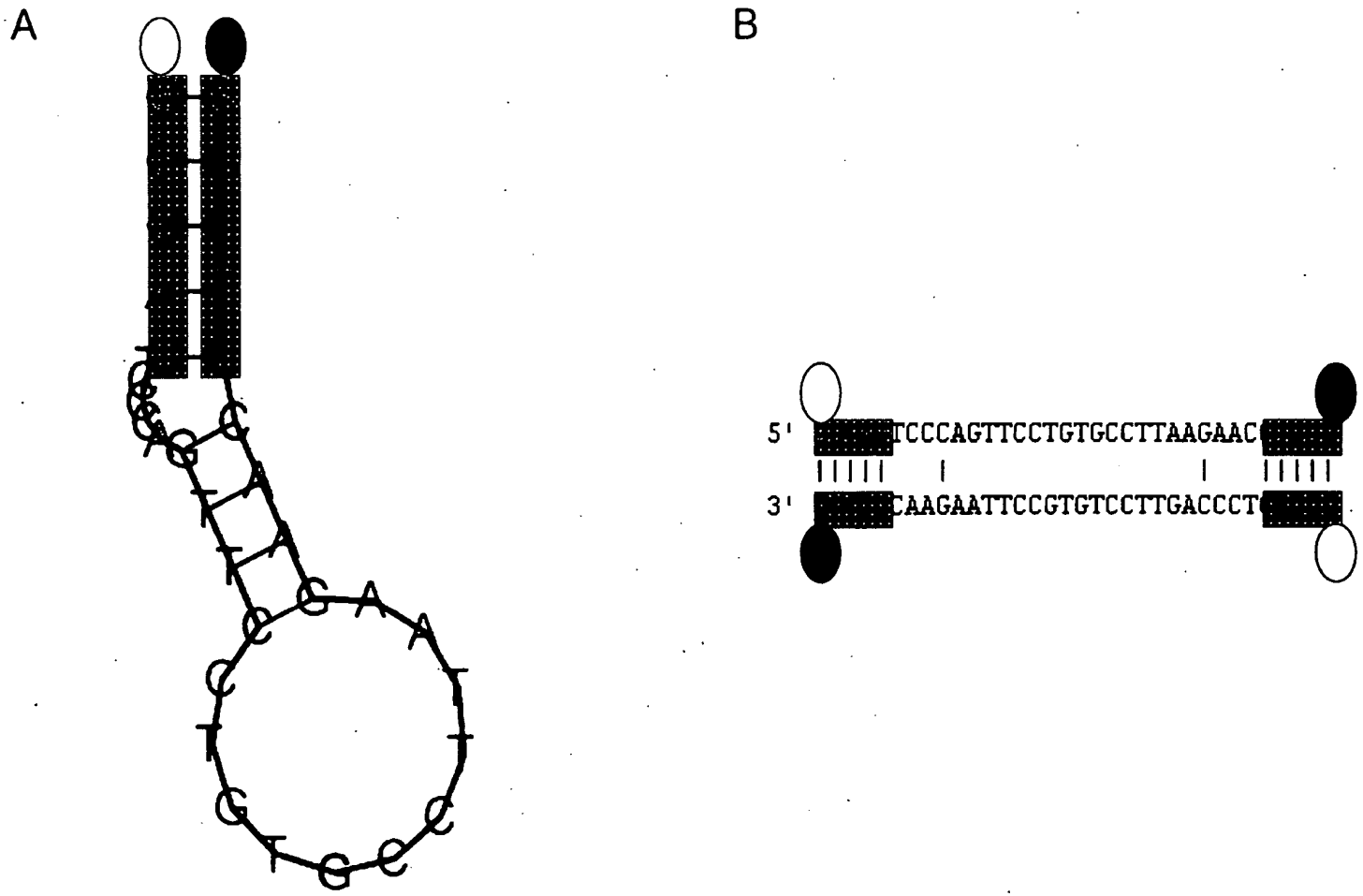


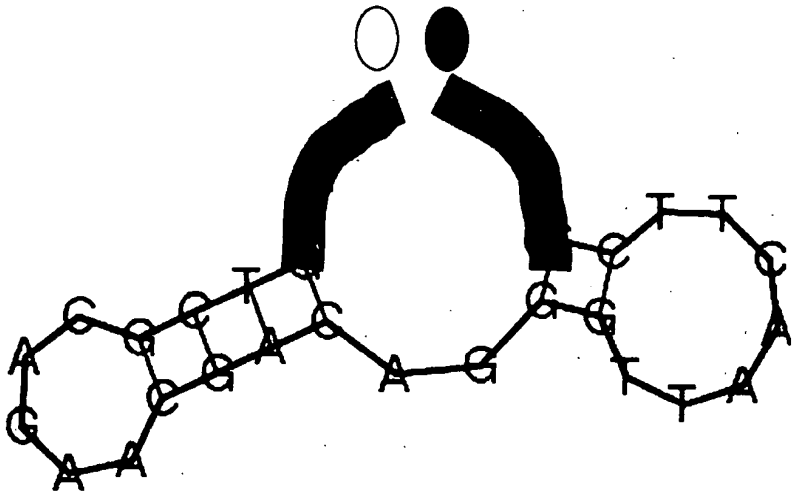
FIG. 7



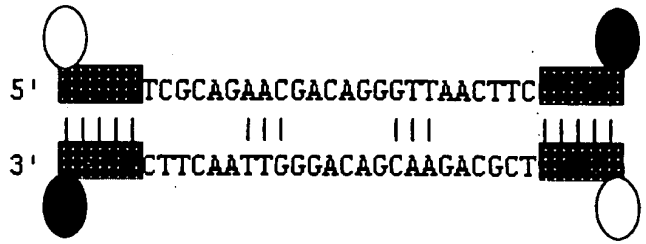
FP1 5' GCGAG TCCCAGTTCCTGTGCCTTAAGAACCTCGC 3'

FIG. 8

A

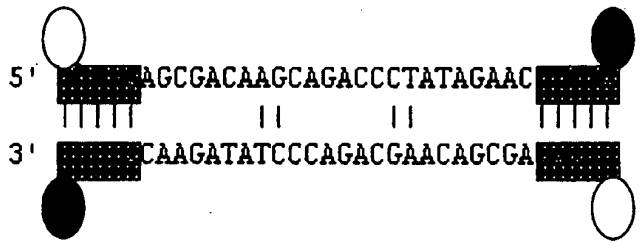


B



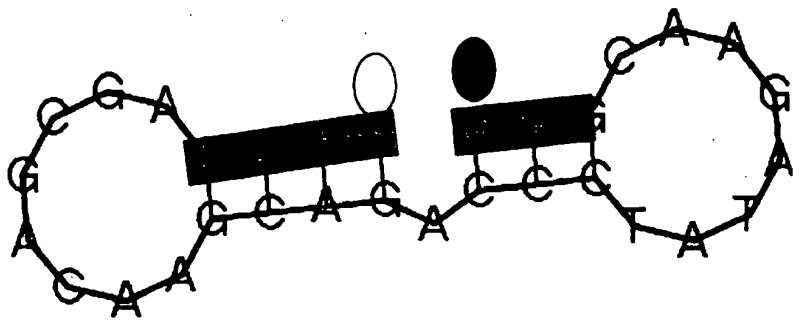
FP2 5' GCGAG TCGCAGAACGACAGGGTTAACTTCCTCGC 3'

FIG. 9



FP3 5' <u>GCGAG</u> AGCGACAAGCAGACCCTATAGAAC <u>CTCGC</u> 3'

FIG. 10



FP4 5' CTGC AGCGACAAGCAGACCCTATAGAAC GGG 3'

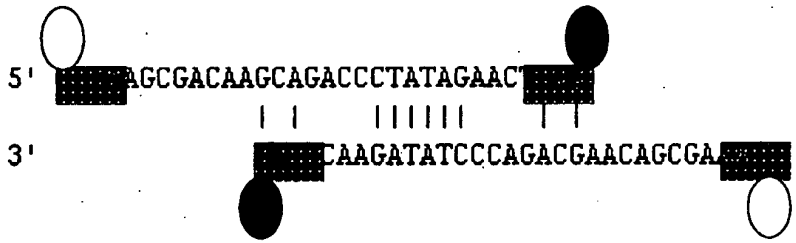
FIG. 11

FP5 5' ggaa TCCCAGTTCCTGTGCCTTAAGAAC ttcc 3'

FIG. 12

FP6 5' ggaa TCGCAGAACGACAGGGTTAACTTC ttcc 3'

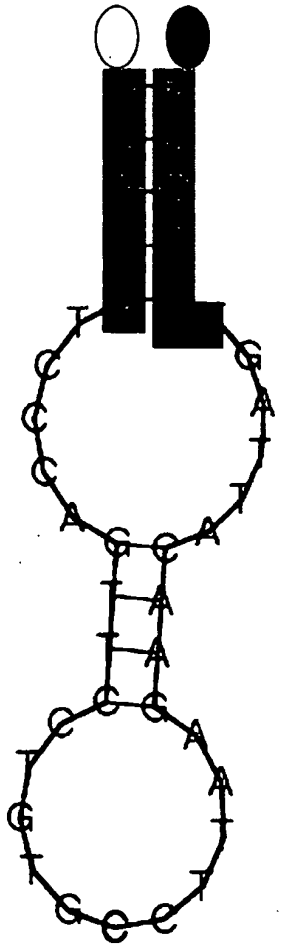
FIG. 13



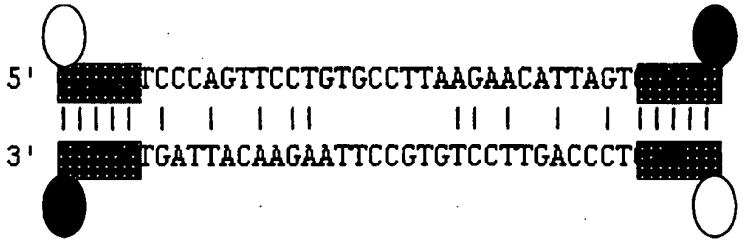
FP7 5' ggaa AGCGACAAGCAGACCCTATAGAAC ttcc 3'

FIG. 14

A



B



FP8 5' GCGAG TCCCAGTTCCTGTGCCTTAAGAACattagtCTCGC 3'

FIG. 15

FP9 5' ggaa TCCCAGTTCCTGTGCCTTAAGAACattagtttcc 3'

FIG. 16

FP10 5' gtgat TCCCAGTTCCTGTGCCTTAAGAAC atcac 3'

FIG. 17

FP11 5' ggac TCCCAGTTCCTGTGCCTTAAGAAC gtcc 3'

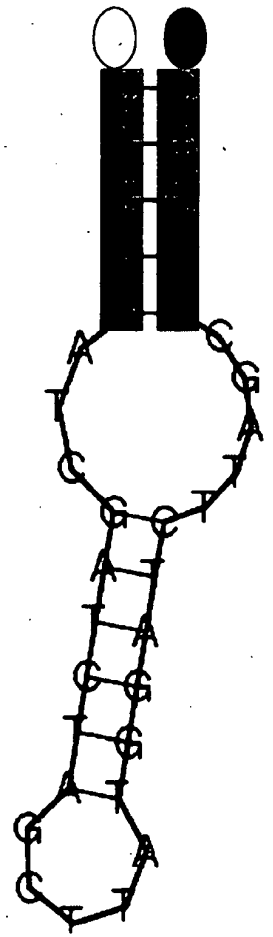
FIG. 18

FP12 5' ggg TCCCAGTTCCTGTGCCTTAAGAAC ccc 3'

FIG. 19

FP13 5' gcg TCCCAGTTCCTGTGCCTTAAGAAC cgc 3'

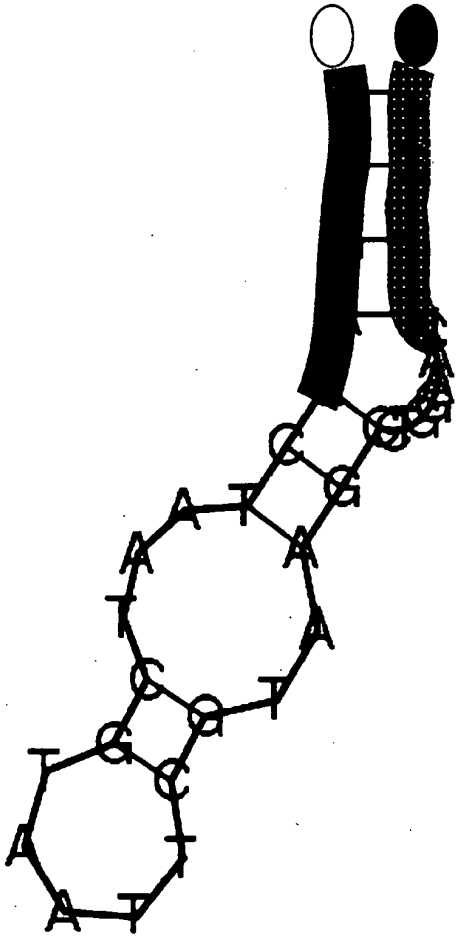
FIG. 20



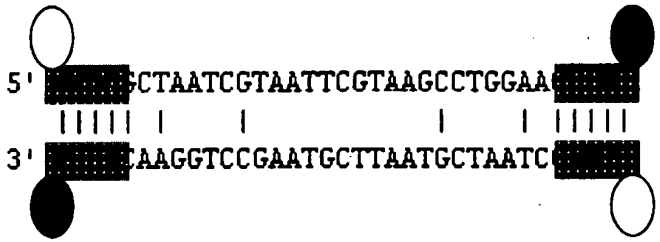
FP14 5' gcgag ATCGATCTAGCTTATGGATCTTAGC CTCGC 3'

FIG. 21

A



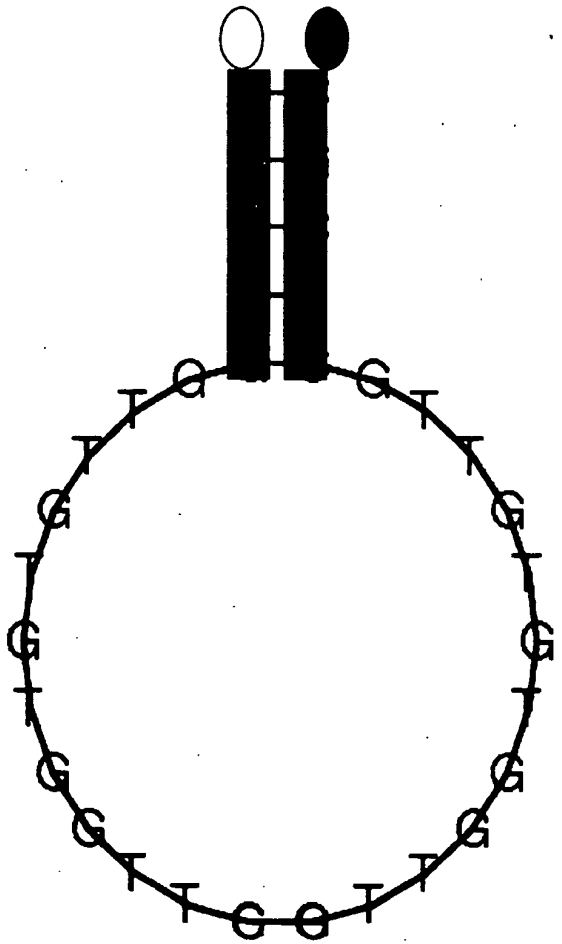
B



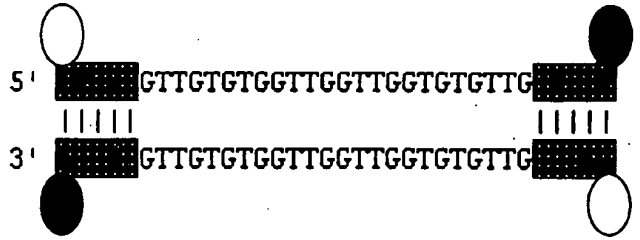
FP15 5' gcgag CTAATCGTAATTCGTAAGCCTGGAA CTCGC 3'

FIG. 22

A



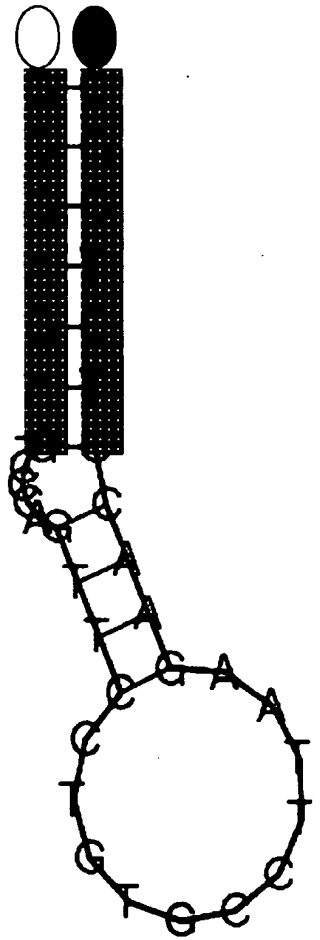
B



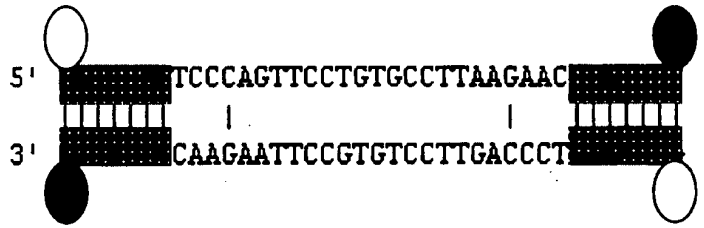
FP16 5' gcgag GTTGTGTGGTTGGTTGGTGTGTTG ctcgc 3'

FIG. 23

A



B



FP17 5' GCGAGAG TCCCAGTTCCTGTGCCTTAAGAAC CTCTCGC 3'

FIG. 24

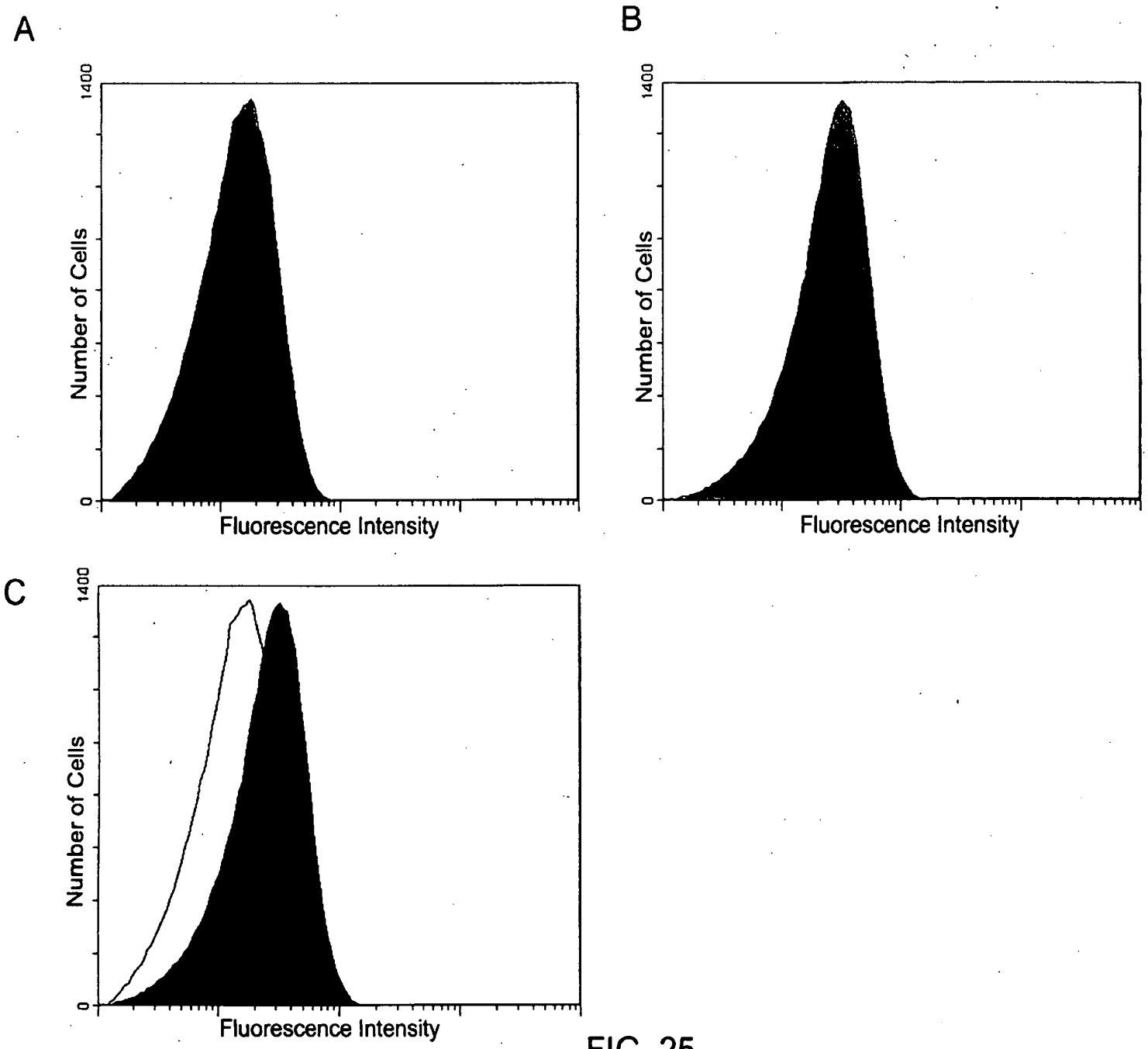


FIG. 25

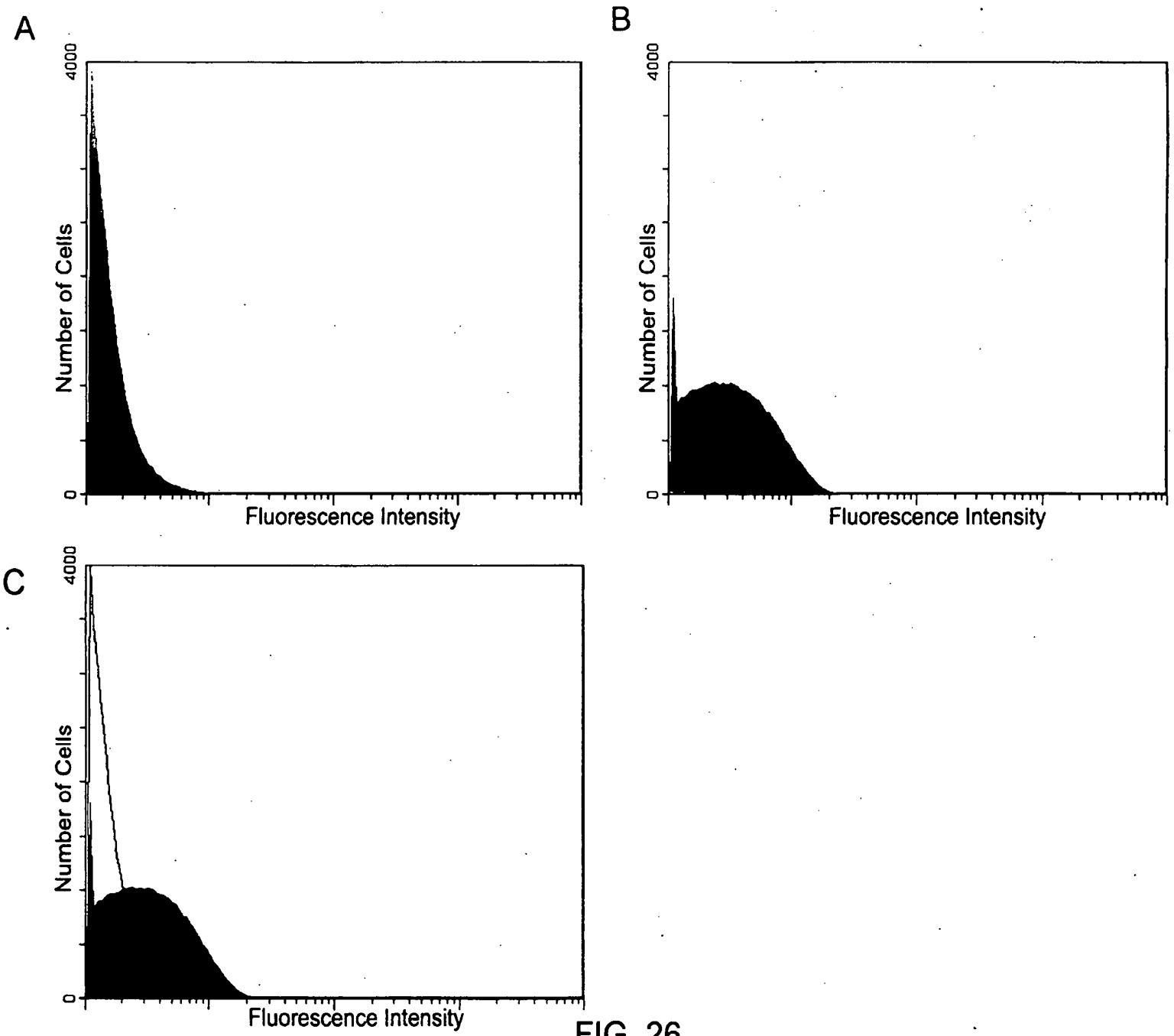

A
4000
0
Number of Cells
Fluorescence Intensity
B
4000
0
Number of Cells
Fluorescence Intensity
C
4000
0
Number of Cells
Fluorescence Intensity

FIG. 26

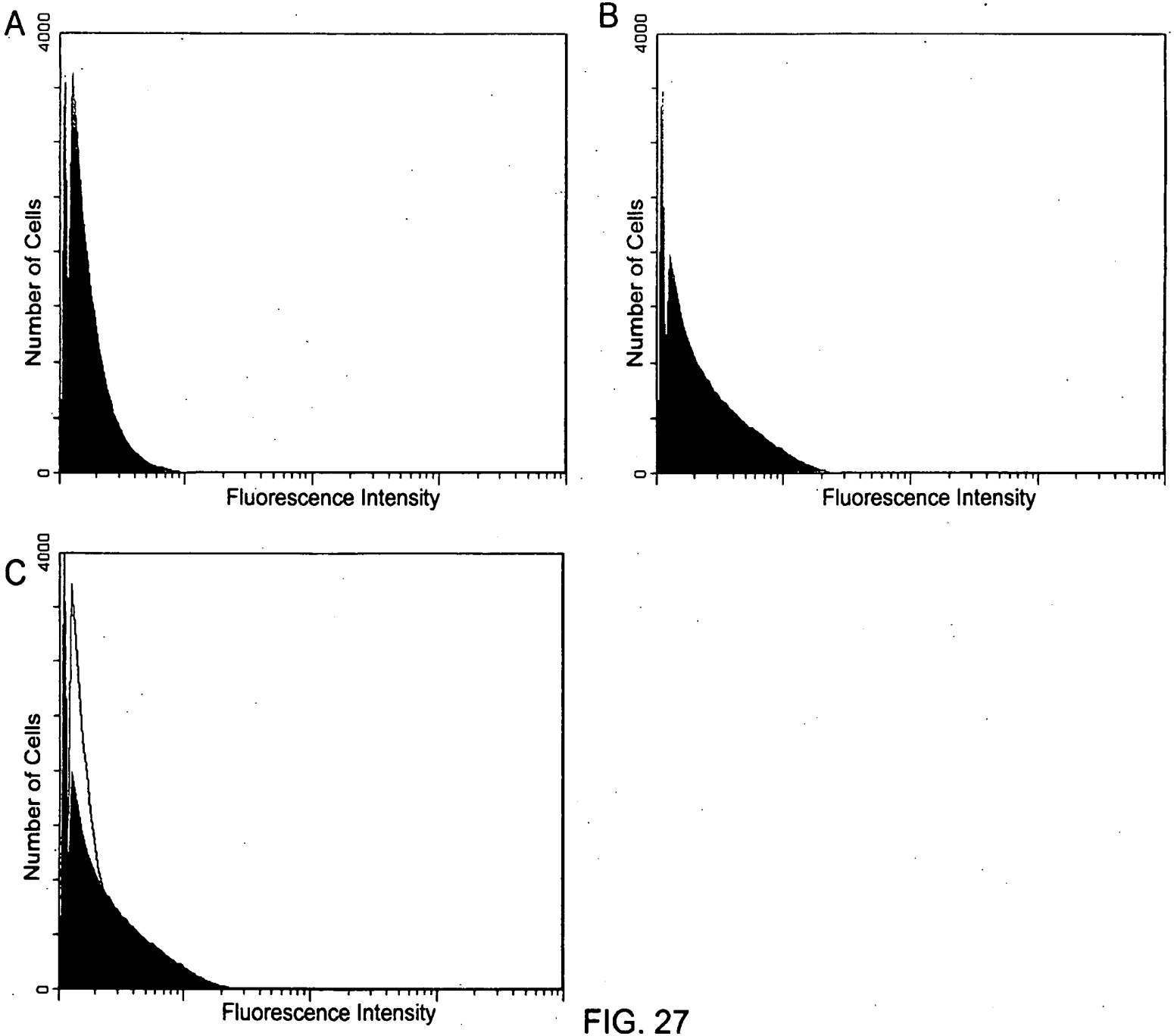

A
4000
0
Number of Cells
Fluorescence Intensity
B
4000
0
Number of Cells
Fluorescence Intensity
C
4000
0
Number of Cells
Fluorescence Intensity

FIG. 27

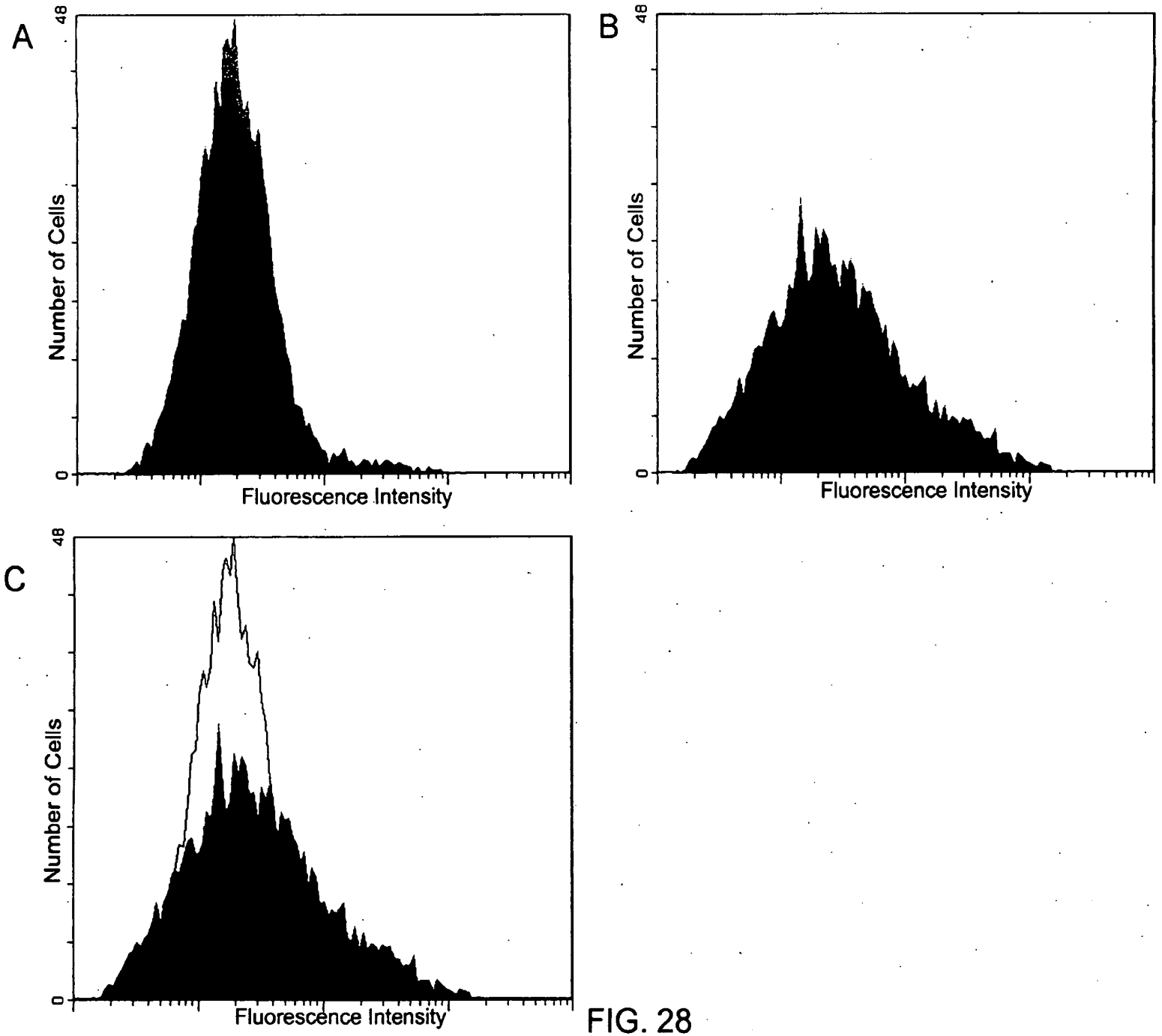

A
48
Number of Cells
0
Fluorescence Intensity
B
48
Number of Cells
0
Fluorescence Intensity
C
48
Number of Cells
0
Fluorescence Intensity

FIG. 28

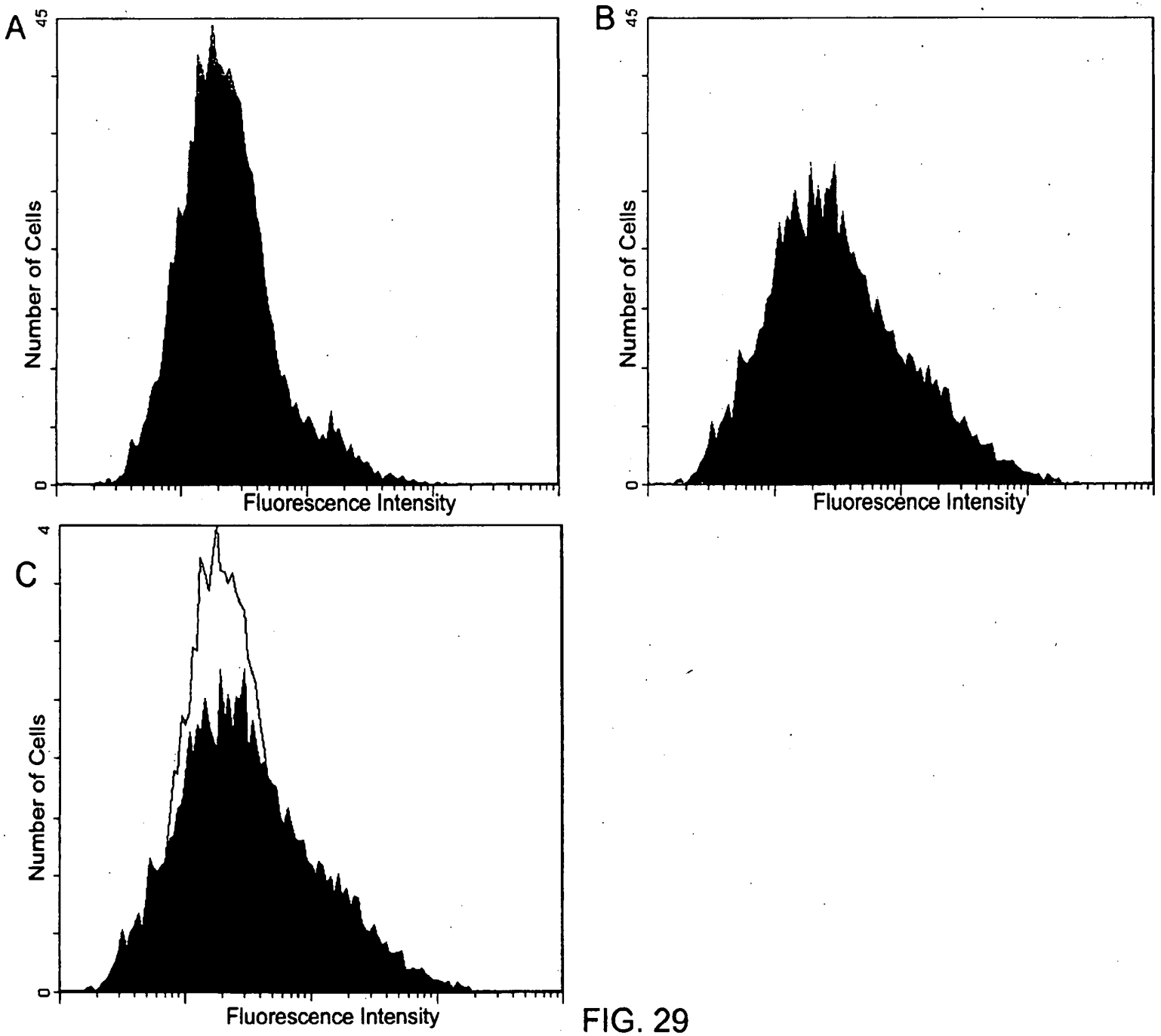

A
45
0
Number of Cells
Fluorescence Intensity
B
45
0
Number of Cells
Fluorescence Intensity
C
4
0
Number of Cells
Fluorescence Intensity

FIG. 29

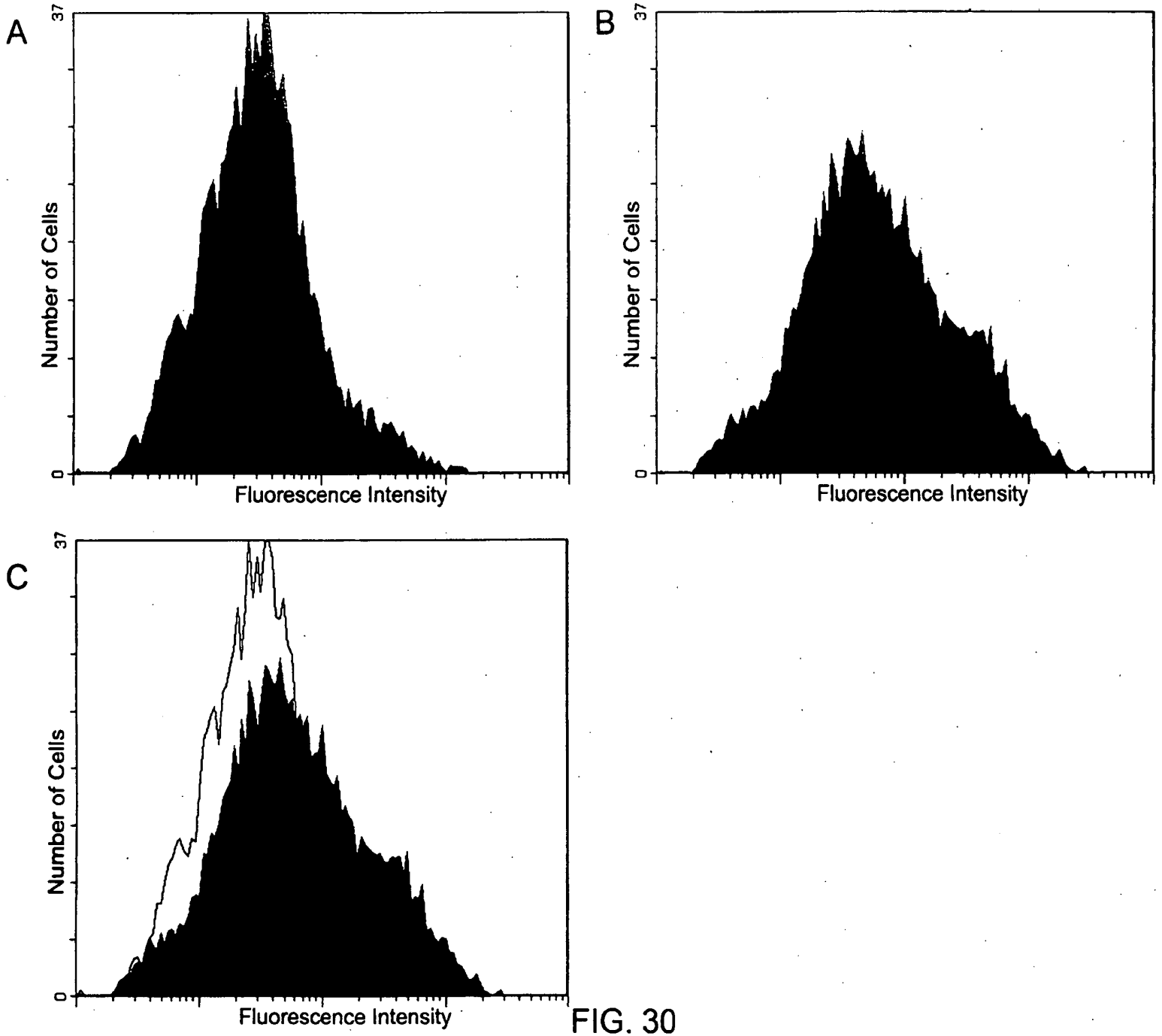

A
37
0
Number of Cells
Fluorescence Intensity
B
37
0
Number of Cells
Fluorescence Intensity
C
37
0
Number of Cells
Fluorescence Intensity

FIG. 30

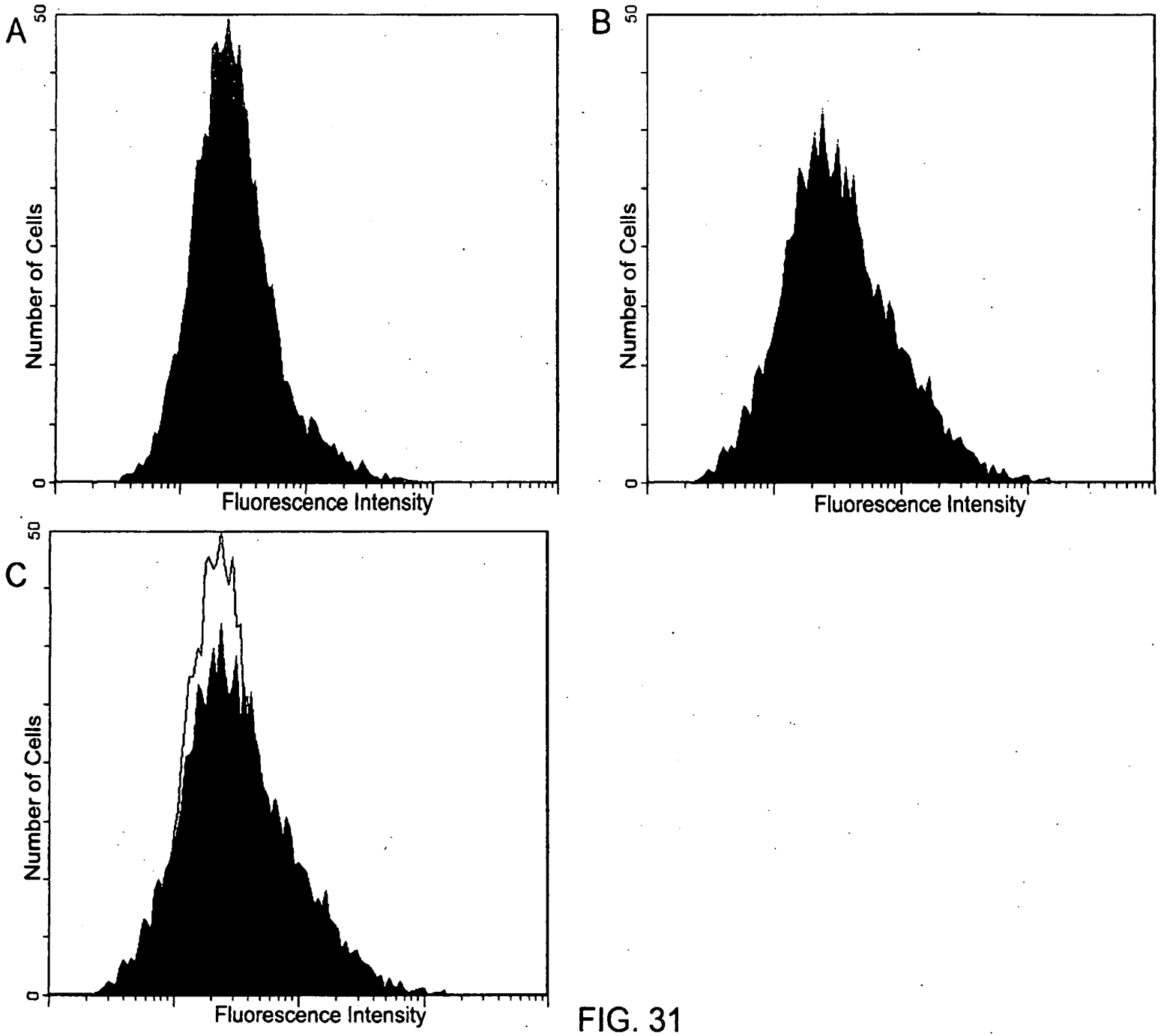

A
50
0
Number of Cells
Fluorescence Intensity
B
50
0
Number of Cells
Fluorescence Intensity
C
50
0
Number of Cells
Fluorescence Intensity

FIG. 31

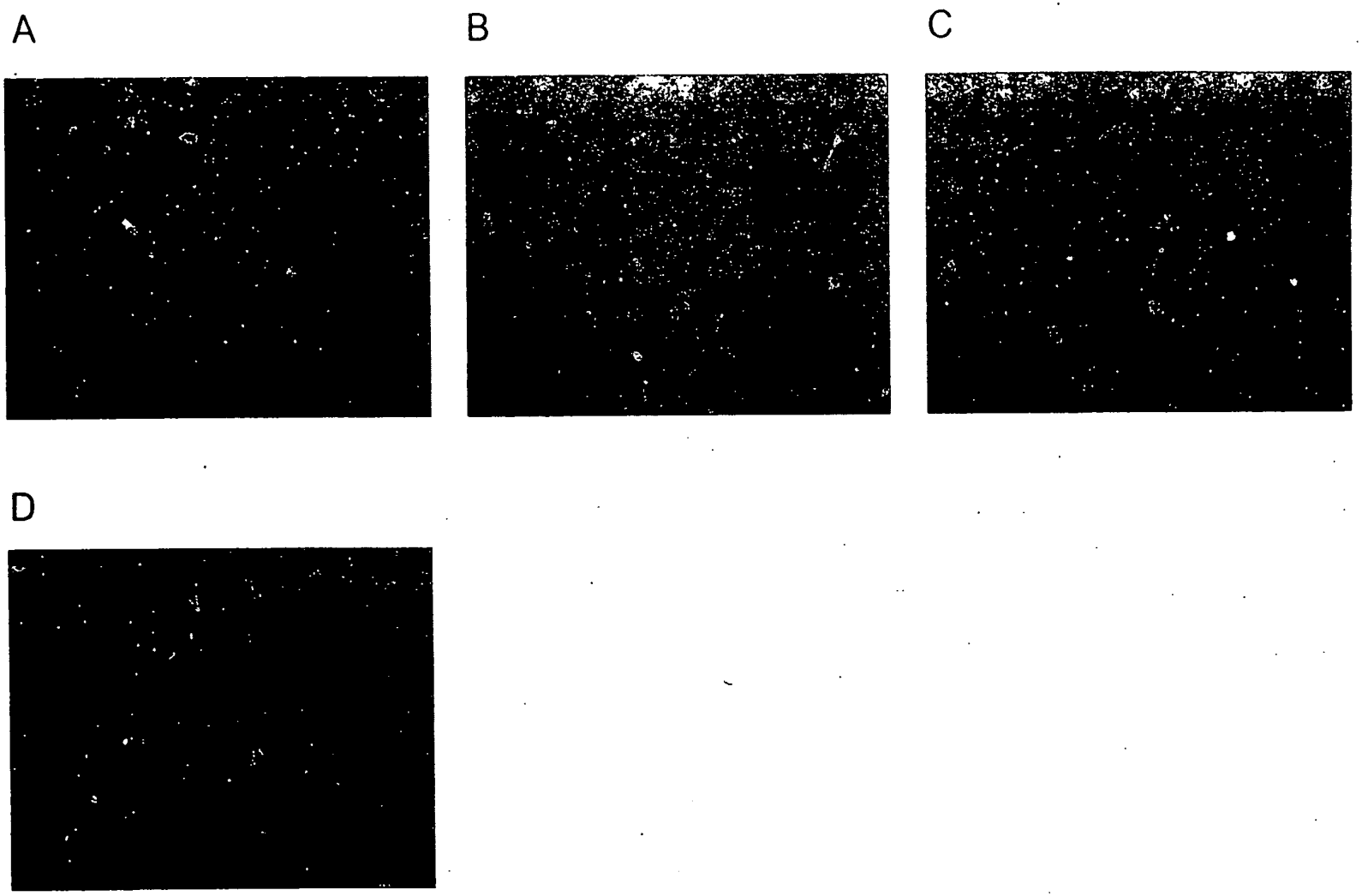


FIG. 32

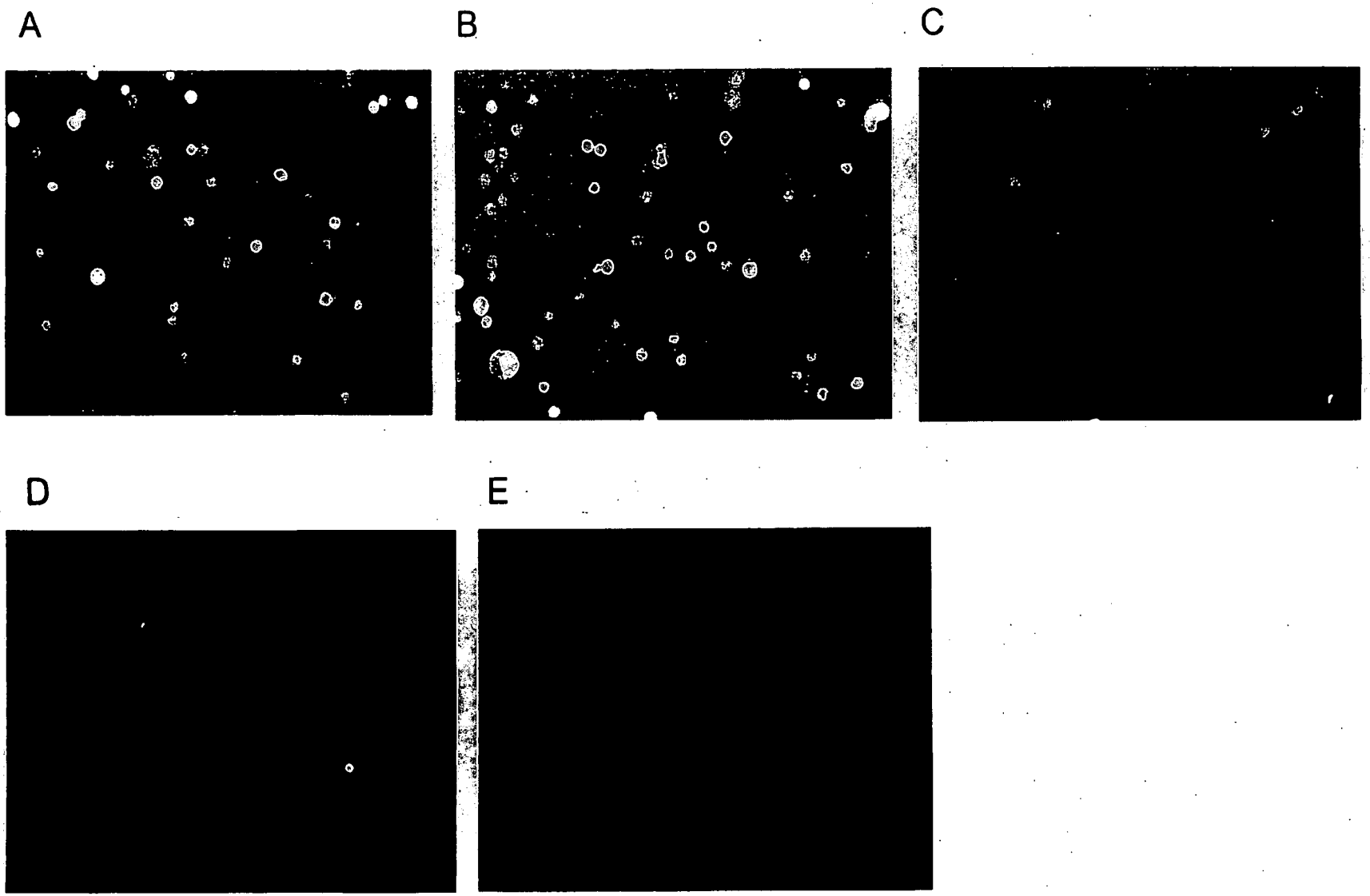


FIG. 33

A

B

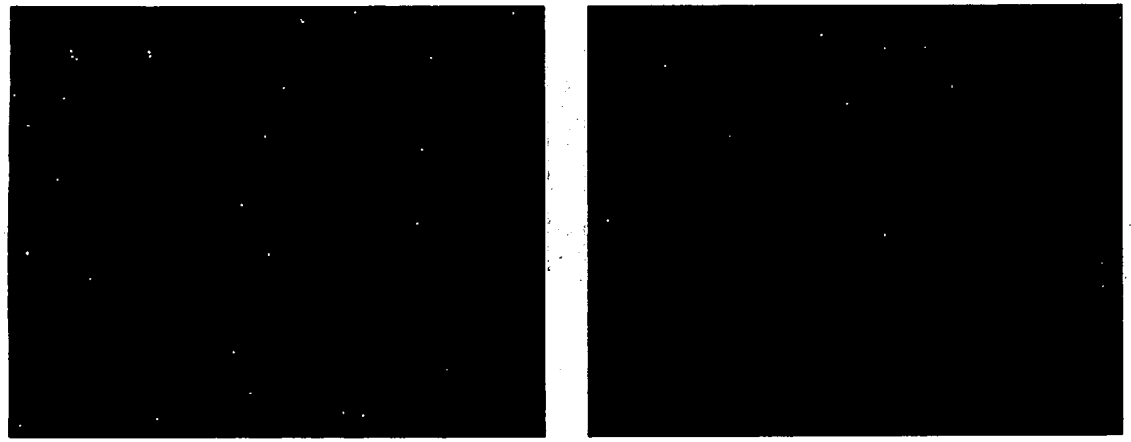


FIG. 34

A

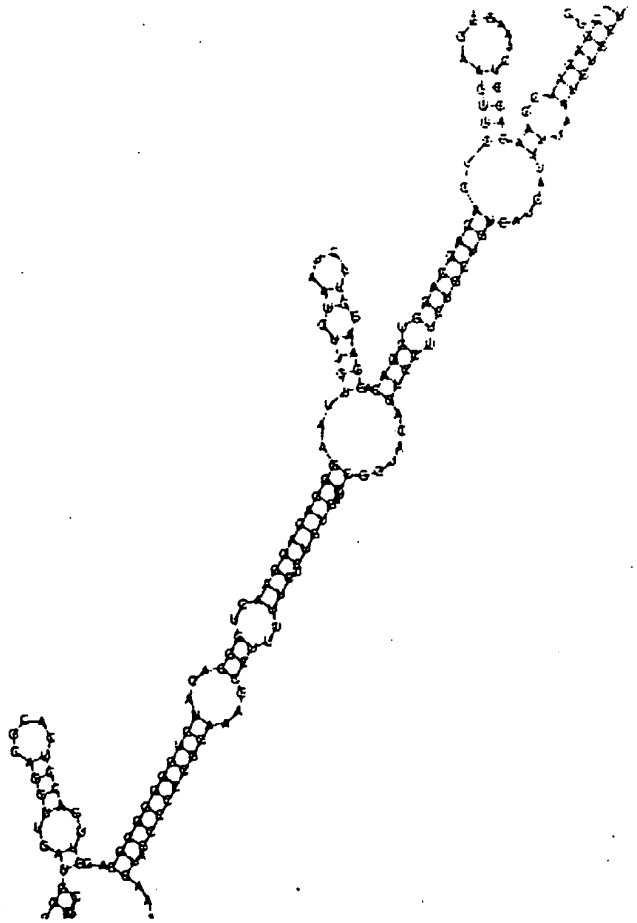
ACCTTGCCAAAATCCTGCACATCGAAGAGGTCAAAGGCTTCGAAGAGCTCGCTCCGCTTGAGGCCGAACTTCTCACA
GCAGGTGGACAGGAAGGTTCTAATGTTCTTAAGGCACAGGAACTGGGACATCTGGGGGCGCCAGGTTGACCTCACGCA
GGTTGATGGCATGGGGTAGCAGGTTGTTAAGCAGCTGACACAGAAGGACACCATCCCGGAGGGCCTGGGCCAGTTCA
CACACCTGAGCCCCATCCCAGGTCACGCGGTGGCTGGGCGGCAGCACCCGGCACTGGATGAGCCAGTGGGTGCATTG
GCGCCACAGCTCCATGGCTACCGCCGGTGGGGAAGGGCATGATCCCCCTGTTCTGGGCGATCGGGGTCCAGGACAGA
GCAGACAGGGTGTAGATGCCTCCACCGTGTCGTCGATCTGGTCGGACAGGCCACTGTAGATGTCTTCATCACCTACA
CTCTCCTCCTCATGAGGTCCTCATAGATCTCGTCGCCTTCCGCCTCCTCATTCTCCACGCAGTCATACAGGTCCTCAT
CCTCTCCCGCAGGCAGCAGCAGCGCTTGTCATACTCTGTCATCTTGGGCGGCATGGACACGGGCTCCGAGCGCGCAG
GGGCTTCAAGAAATGCTGCTGGATGGAGCCCAGCGTGTCAGTGTACTTCTCCTCCGTCTGCTGGATGTATGAACACGA
AGCAGGTCCTCAATGTTGATAAAGATGATCTCAATGTCTTGAGGTTTCAGGAACCGTTGATGAAGACCTGGTAGAGAT
TGGCTGCGCCAGGGGTGCCCAGGGCTTCCTTCATCTCCTTTAGGAAGTGACCAGGTGTTTGCTGGCTGACTCCACCTG
GCTGCAGTAGCGGCCATAGACGAGGAACCTCTCCTTGTATTTCCGTTGTTGGCTCTCTGAGAACATTCCTCCAGCTTC
ATCTGCACGTCCTCCCGGGCTGCGGCCACACGGTCTGGAGAAGGAGGTGATATTTGAGAACTCGCTGCATAGGCACC
ATCAGCAGGTCCCGCAGGGTGAACCTCCCCTCATGGCATCCAGGGCCAGCCGCAGGTTCTCCTTCTCCATCGCCTCC
TGCGTGTGTTTCACCAGCTCAGCTGGAAATTGGTGATCTGTCGCAGTGTCTCGTTGTCTCGCTTGACCTCGTTCACGC
ACTGAGCCAGGTGATCTTGAGTTCCCCGTCGATCTTGGGCCGGCCATAGTGAGCCAGAGACTGGTCCAGGTTCTCAA
TGGACTACAGATGAGTAGAGCTTTGTCGAGCAGGAAGGCATACCTGTCCATCTTGGAGCGCCGTTCCACCGAGGTGA
GTCATCCCGAACCTGGAAGCTGTGCAGGTTTACAAAGTCCTTGAGGTCATAGGAGTCTCCCCTGCGCTGCCCTGGGC
ACCTTGGTCCTCGATCAGGAGGAACATGTGGCTCCACTTCTTGTTGTCTCGGTCTCCTGAATGGAGATGGCCATCTCA
AACTGCTCCATCCACTTCTTCTTCAATTCTCTTGTCTTGAAGAACAGCTCATAGATGTGGTCTCCTCAAAGGAGAACA
TCTGGAAGTCATGCCCGTTGGCGGTGGCATTCTCCGGATAGATGTTGCCCGGCACCGATGGCAGCGGTAGCCCTGAT
AGAAGGTACCTCTAAGCAGCATCTGACAGGCCTTGCAGTAGTTCCTGGGAAATCTTGCCCATGTCGGCCACATGGAG
GGACCCTCCCCAGACACTCCTTGTGTGCAGATCCATCTTGGGCAGACCCAGCTCATTCCTCTTTTTGTCCTGAGCCCT
GCGATGTAGTTTGTCCTTCTTCAAGGGTTGAGCCGTAGAAAGGGTCCAATGGCTCCAGGGGGTGGAGGAAGCCCGTA
GTATTCCTGAAACACCTAGTAGATGTATTTCTGCCCTCCCACCAGTTCTGTTCAGCCTCAGCCTTCGTGAGCTCCACA
ATGTCTCCACAGACAGGTCCTGAGGAGGGCCATGGACATAGGGCTTCACCCTGTTACAAGGAAACCAGCCAATTTCA
TCCCGTCCGAGCGGTTGGCCAGGATGCTCTCTGCCCCTGCCCGCTCCATGGGGCCTGCGTACCAGAGATGACCTCGA
CGTTATATTTAATGCTGATGGCAAATTCTGCTGCATCCTTCACCCTCTGCCGCACCAAGAAAGTAGCCCCCGAAAGC
CTTTTTCTCTGTGATCCGGTACAGTCCTTCTGCTGTCATGATTTTAATGTGCTTGACAAGGTGGTGTCCAGAGACTTGA
AGCAATCCTTTAGAGAGTTCTGCTGGTAAAACTCCACCAGCTCCGTACAAAATACTTTGTGCTTCCCACTGCTGGCCT
GCTGATGGTTCTCTTTTCAGGCTCCTTGAAGGGGAACTGATGTCACCCTCCTTGAGCGACAGCTCTGATCGGTCTCGG
GCGCAGAAGTCATAGCGGGCTTTGGCTGTGCCCAGCCAACCCGGCCATAGATCTCCCCTCGCCACCAGCCTTGCTGT
CCCTTCTTGTTAAGGATCTTGATGCGTCTCTCTGCCAAGGCACCAGGGCTCAGCAGTATTCAGAATAATCTTCCTCCA
CGTAGTTGGCAGGGAAACCCGCCGGGAGCTGTCACAGCCCCTGGCTCCGCTGCCTGCCCACGCCGGGCTCAGAGCCT
GGAGTTTCTCATGTCAGGGCACATCCCGGGAGCACCGCCAGCTGGACGCTGCGCTCCTCCAGTCCATCCCAAAGTCT
CCCCTTGAGCCCCGTCTGCTTGAGGGGACCCAAGAGGAAATCGGGGTGTTATAAATTAAC

B

GCAGGTGGACAGGAAGGTTCTAATGTTCTTAAGGCACAGGAACTGGGACATCTGGGCCCCGGAAAGCCTTTTTCTCTGT
GATCCGGTACAGTCCTTCTGC

FIG. 35

A



B

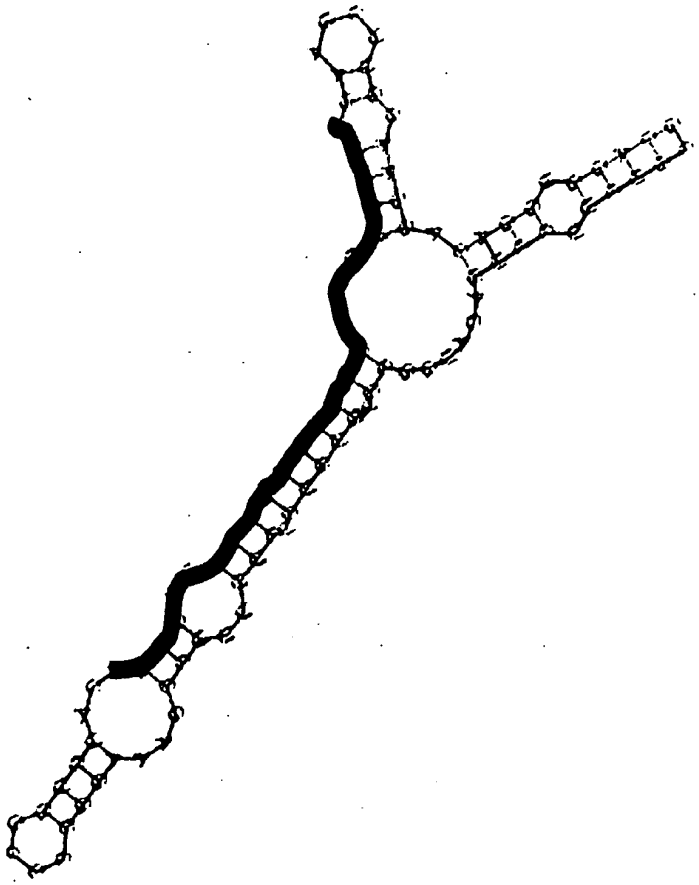


FIG. 36

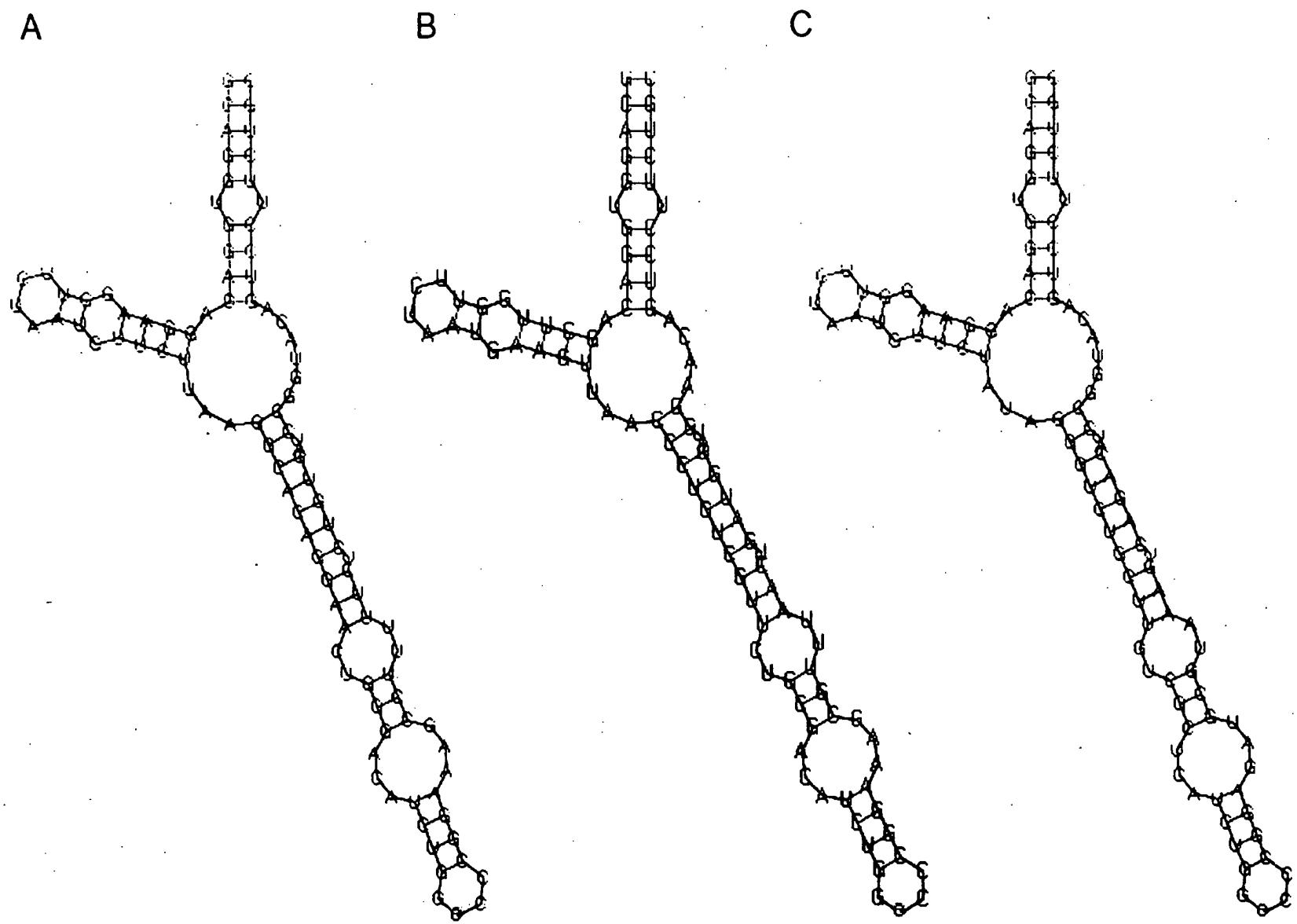


FIG. 37

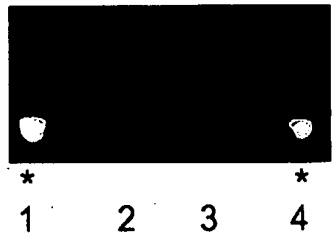


FIG. 38

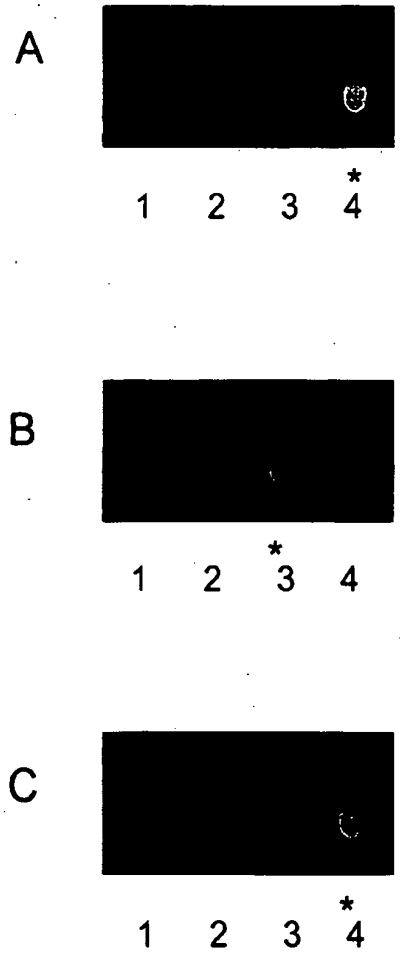


FIG. 39

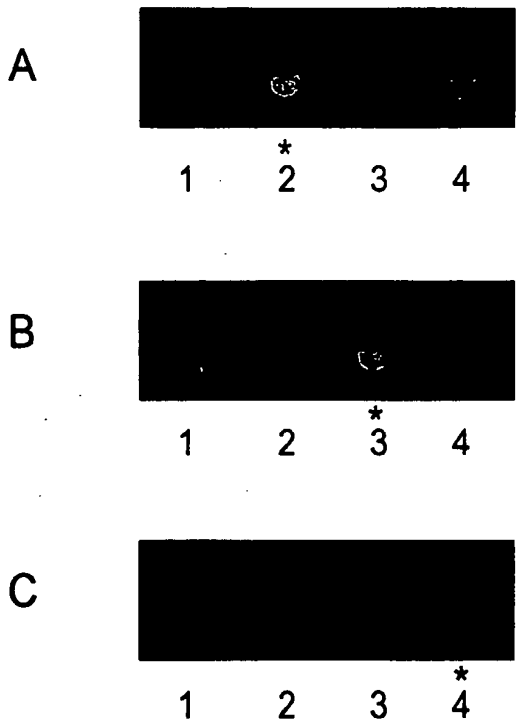
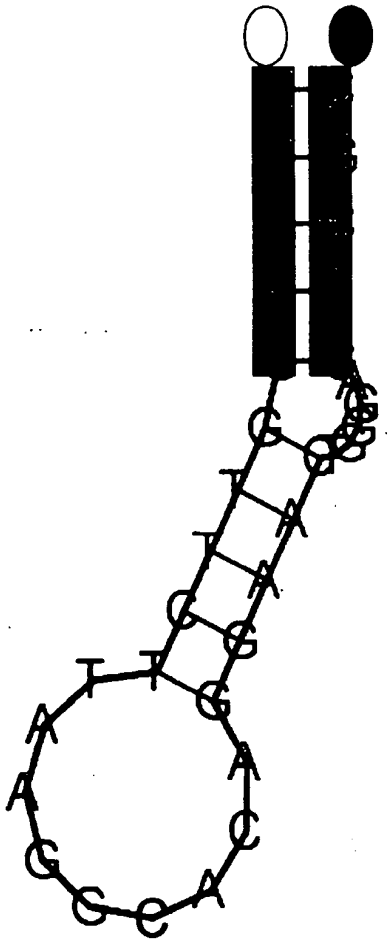
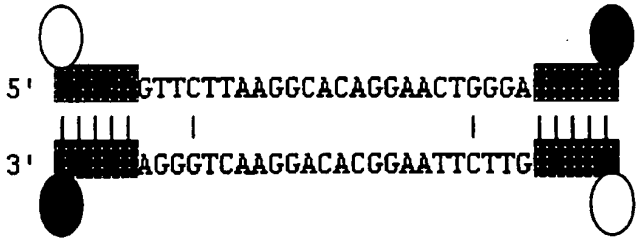


FIG. 40

A



B



FP18 5' GCGAG GTTCTTAAGGCACAGGAACTGGGA CTCGC 3'

FIG. 41

A

GCAGGTGGACAGGAAGGTTCTAAT**GTTCTTAAGGCACA**
GGAACTGGGACATCTGGGCCCGGAAAGCCTTTTTCTCT
GTGATCCGGTACAGTCCTTCTGC

B

GCAGGTGGACAGCTTGGTTCTAAT**GAAGTTAACCCTGT**
CGTTCTGCGACATCTGGGCCCGGAAAGCGTTTAACUGA
TGGAUGGAACAGTCCTTCTGC

C

GCAGGTGGACAGGAAGGTTCTAAT**GTTCTATAGGGTCT**
GCTTGTCGCTCATCTGGGCCCGGAGATGCGT
AAAGTCAGACATCCGGTACAGTCCTTCTGC

D

1	GTTCTTAAGGCACAGGAACTGGGA	
2	GAAGTTAACCCTGTCGTTCTGCGA	50% different from target 1; 66% different from target 3
3	GTTCTATAGGGTCTGCTTGTCGCT	50% different from target 1

FIG. 42

A start
GCAGGTGGACAGGAAGGTTCTAATGTTCTTAAGGCACAGGAACTGGGACATCTGGGCCCGGAAAGCCTTTTTCTCTGTGATCCGGTACAGTCCTTC
TGC

B 1
GCAGGTGGACAGCTTGGTTCTAATGAAGTTAACCCTGTCGTTCTGCGACATCTGGGCCCGGAAAGCGTTTAACTGACAGATGGGGTACAGTCCTTC
TGC

C 2
GCAGGTGGACAGCTTGGTTCTAATGAAGTTAACCCTGTCGTTCTGCGACATCTGGGCCCGGAAAGCGTTTAAC[TGACA]GATGGGGTACAGTCCT
TCTGC

D 3
GCAGGTGGACAGCTTGGTTCTAATGAAGTTAACCCTGTCGTTCTGCGACATCTGGGCCCGGAAAGCGTTTAAC[[T]]GATGGGGTACAGTCCTTC
TGC

E 4
GCAGGTGGACAGCTTGGTTCTAATGAAGTTAACCCTGTCGTTCTGCGACATCTGGGCCCGGAAAGCGTTTAACTGATGG[[AT]]GGTACAGTCCT
TCTGC

F 5
GCAGGTGGACAGCTTGGTTCTAATGAAGTTAACCCTGTCGTTCTGCGACATCTGGGCCCGGAAAGCGTTTAACTGATGGATGGAACAGTCCTTCTG
C

G modifications of sequence compared to previous sequence:
1 underlined = change in bases
2 [single bracket] = bases cut out
3 [[double bracket]] = bases added
4 [[double bracket]] = bases added
5 underlined = change in bases

FIG. 43

A

start

B GCAGGTGGACAGGAAGGTTCTAATGTTCTTAAGGCACAGGAACTGGGACATCTGGGCCCGGAAAGCCTTTTTCTCTGTGATCCGGTACAGTCCTTCTGC

1

GCAGGTGGACAGGAAGGTTCTAATGTTCTATAGGGTCTGCTTGTCGCTCATCTGGGCCCGGAAAGCCTTTAAGTCAGACATCCGGTACAGTCCTTCTGC

C

2

GCAGGTGGACAGGAAGGTTCTAATGTTCTATAGGGTCTGCTTGTCGCTCATCTGGGCCCGGAAAGCCTTAAAGTCAGACATCCGGTACAGTCCTTCTGC

D

3

GCAGGTGGACAGGAAGGTTCTAATGTTCTATAGGGTCTGCTTGTCGCTCATCTGGGCCCGGAAAGCGTTAAAGTCAGACATCCGGTACAGTCCTTCTGC

E

4

GCAGGTGGACAGGAAGGTTCTAATGTTCTATAGGGTCTGCTTGTCGCTCATCTGGGCCCGGAATGCGTTAAAGTCAGACATCCGGTACAGTCCTTCTGC

5

GCAGGTGGACAGGAAGGTTCTAATGTTCTATAGGGTCTGCTTGTCGCTCATCTGGGCCCGGAATGCGT[T]

F AAAGTCAGACATCCGGTACAGTCCTTCTGC

6

GCAGGTGGACAGGAAGGTTCTAATGTTCTATAGGGTCTGCTTGTCGCTCATCTGGGCCCGGA[[G]]ATGCGT

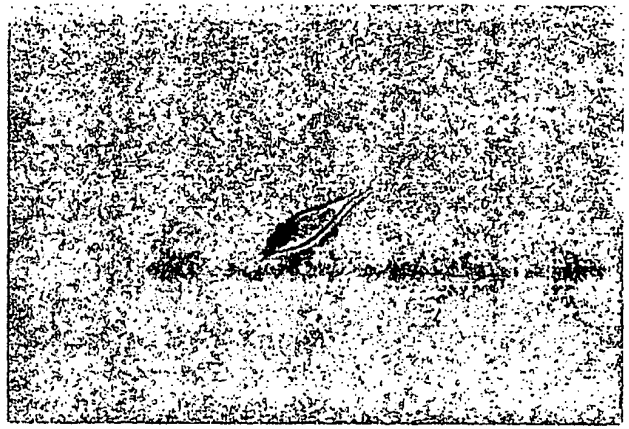
G AAAGTCAGACATCCGGTACAGTCCTTCTGC

1 underlined = change in bases
2 underlined = change in bases
H 3 underlined = change in bases
4 underlined = change in bases
5 [single bracket] = bases cut out
6 [[double bracket]] = bases added

FIG. 44

A

B



C

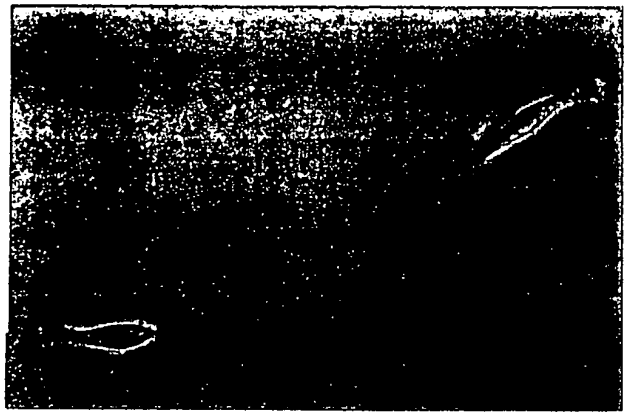


FIG. 45

INVENTOR INFORMATION

Inventor One Given Name:: Kambiz
Family Name:: Shekdar
Postal Address Line One:: 29 Fifth Avenue, Apt. 6A
City:: New York
State or Province:: New York
Country:: U.S.A.
Postal or Zip Code:: 10003
City of Residence:: New York
State or Province of Residence:: New York
Country of Residence:: U.S.A.
Citizenship Country:: U.S.A.
Inventor Two Given Name:: Dennis J
Family Name:: Sawchuk
Postal Address Line One:: 77 Park Avenue, Apt. 1007
City:: Hoboken
State or Province:: New Jersey
Country:: U.S.A.
Postal or Zip Code:: 07030
City of Residence:: Hoboken
State or Province of Residence:: New Jersey
Country of Residence:: U.S.A.
Citizenship Country:: Canada
Inventor Three Given Name:: Jason M
Family Name:: Montez
Postal Address Line One:: 1500 Washington Street, Apt. 7E
City:: Hoboken
State or Province:: New Jersey
Country:: U.S.A.
Postal or Zip Code:: 07030
City of Residence:: Hoboken
State or Province of Residence:: New Jersey
Country of Residence:: U.S.A.
Citizenship Country:: U.S.A.

CORRESPONDENCE INFORMATION

Correspondence Customer Number:: 1473
Fax One:: 212.596.9090
Electronic Mail One:: jgunnison@fishneave.com

APPLICATION INFORMATION

Title Line One:: ISOLATION OF LIVING CELLS, PREPARATION O
Title Line Two:: F CELL LINES QUANTIFICATION OF PRESELECT
Title Line Three:: ED CELLULAR RIBONUCLEIC ACID SEQUENCES B
Title Line Four:: ASED ON SIGNALING PROBES
Total Drawing Sheets:: 45

Formal Drawings?:: No
Application Type:: Provisional
Docket Number:: CHROMO/3PROV
Secrecy Order in Parent Appl.?:: No

REPRESENTATIVE INFORMATION

Representative Customer Number:: 1473
Registration Number One:: 27794
Registration Number Two:: 30709
Registration Number Three:: 38479
Registration Number Four:: 42800
Registration Number Five:: 45141

Source:: PrintEFS Version 1.0.1